

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МІСЬКОГО ГОСПОДАРСТВА імені О. М. БЕКЕТОВА

К. А. Мамонов
І. С. Творошенко

МАТЕМАТИЧНІ МЕТОДИ І МОДЕЛІ В ОЦІНЦІ НЕРУХОМОСТІ

НАВЧАЛЬНИЙ ПОСІБНИК

ХАРКІВ
ХНУМГ
2014

УДК [519.87:347.235](075)

ББК 65.225.1в631я73-6

М22

Автори:

Мамонов Костянтин Анатолійович, доктор економічних наук, завідувач кафедри геоінформаційних систем, оцінки землі та нерухомого майна ХНУМГ ім. О. М. Бекетова;

Творошенко Ірина Сергіївна, кандидат технічних наук, доцент кафедри геоінформаційних систем, оцінки землі та нерухомого майна ХНУМГ ім. О. М. Бекетова.

Рецензент:

Шостак Ігор Володимирович, доктор технічних наук, професор кафедри інженерії програмного забезпечення Національного аерокосмічного університету імені М. Є. Жуковського «ХАІ»;

Дорошенко Володимир Олексійович, доктор фізико-математичних наук, декан факультету прикладної математики та менеджменту Харківського національного університету радіоелектроніки;

Ачкасов Анатолій Єгорович, доктор економічних наук, завідувач кафедри економіки підприємств міського господарства Харківського національного університету міського господарства імені О. М. Бекетова.

Рекомендовано на засіданні

*Вченої ради ХНУМГ ім. О. М. Бекетова,
протокол № 10 від 30.05.2014 р.*

Мамонов К. А.

М22 Математичні методи і моделі в оцінці нерухомості: навч. посіб. / К. А. Мамонов, І. С. Творошенко; Харк. нац. ун-т міськ. госп-ва ім. О. М. Бекетова. – Х. : ХНУМГ, 2014. – 212 с.

У навчальному посібнику розглядаються основні математичні методи і моделі, які використовуються під час оцінки нерухомості. Викладаються питання щодо організації оптимізаційного моделювання, вирішення задач лінійного та нелінійного програмування, невизначеності та ризику, економетричного моделювання, методичних засобів нечіткого логічного виведення під час оцінки нерухомості.

Навчальний посібник буде корисним для студентів усіх форм навчання освітньо-кваліфікаційного рівня «магістр» спеціальності 8.08010104 «Оцінка землі та нерухомого майна» та працівникам органів державної влади, місцевого самоврядування, підприємств і установ, які є суб'єктами оціночної діяльності, викладачам і студентам вищих навчальних закладів.

УДК [519.87:347.235](075)

ББК 65.225.1в631я73-6

© К. А. Мамонов, І. С. Творошенко, 2014

© ХНУМГ ім. О. М. Бекетова, 2014

ЗМІСТ

ВСТУП.....	5
РОЗДІЛ 1 ОРГАНІЗАЦІЯ ОПТИМІЗАЦІЙНОГО МОДЕЛЮВАННЯ В ОЦІНЦІ НЕРУХОМОГО МАЙНА.....	7
1.1 Концептуальні аспекти оптимізаційного моделювання в оцінці нерухомого майна.....	7
1.2 Оптимізаційні моделі в оцінці нерухомого майна: напрями формування та особливості застосування.....	32
РОЗДІЛ 2 ЛІНІЙНЕ ТА НЕЛІНІЙНЕ ПРОГРАМУВАННЯ В ОЦІНЦІ НЕРУХОМОГО МАЙНА.....	39
2.1 Задача лінійного програмування та методи її розв’язання.....	39
2.2 Теорія достовірності та аналіз лінійних моделей оптимізаційних задач.....	64
2.3 Цілочислове програмування.....	74
2.4 Нелінійні оптимізаційні моделі в оцінці нерухомого майна.....	84
РОЗДІЛ 3 УПРАВЛІННЯ РИЗИКАМИ ПІД ЧАС ОЦІНКИ НЕРУХОМОГО МАЙНА.....	98
3.1 Поняття, сутність і зміст невизначеності та ризику під час оцінки нерухомого майна.....	98
3.2 Система показників кількісного оцінювання ступеня ризику в оцінці нерухомого майна.....	101
РОЗДІЛ 4 ЕКОНОМЕТРИЧНЕ МОДЕЛЮВАННЯ В ОЦІНЦІ НЕРУХОМОГО МАЙНА.....	113
4.1 Принципи побудови економетричних моделей. Парна лінійна регресія.....	113
4.2 Лінійні моделі множинної регресії.....	130
4.3 Узагальнені економетричні моделі в оцінці нерухомого майна.....	144
4.4 Економетричні моделі динаміки в оцінці нерухомого майна.....	149

РОЗДІЛ 5 МЕТОДИ МОДЕЛЮВАННЯ У ЗАДАЧАХ ОПЕРАТИВНОГО ОЦІНЮВАННЯ СТАНІВ СКЛАДНИХ ОБ’ЄКТІВ НЕРУХОМОГО МАЙНА З ВИКОРИСТАННЯМ НЕЧІТКИХ ІНТЕРВАЛЬНИХ УЯВЛЕНЬ.....	182
5.1 Розробка методу оперативного оцінювання станів складних об’єктів нерухомого майна з використанням нечітких інтервальних уявлень.....	182
5.2 Формалізація і розробка розширеного методу дихотомії для задачі настроювання параметрів функцій належності.....	197
5.3 Програмна реалізація розробки.....	204
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	207

ВСТУП

У сучасних умовах господарювання для забезпечення та зростання ефективності функціонування вітчизняних підприємств, суб'єктів регіонального розвитку необхідною умовою є розробка й реалізація сучасних підходів щодо оцінки нерухомого майна, створення відповідної методологічної основи. В цьому аспекті виникає необхідність використання інструментарію, який органічно поєднує математичні методи для вирішення даних проблем з метою отримання кількісних оцінок і моделей в процесі прийняття ефективних управлінських рішень.

На жаль, за останні роки методологія побудови систем оцінки нерухомого майна вдосконалюється тільки за рахунок створення технологій, які враховують різні аспекти та особливості територій, по якій здійснюється оцінка. При цьому поза рамками залишається комплекс питань, які пов'язані із можливістю застосування математичних методів і моделей в оцінці нерухомості.

Метою проведення дослідження є формування системи знань з методології та інструментарію побудови і використання різних типів оптимізаційних моделей в оцінці нерухомого майна.

Предметом дослідження є методологія та інструментарій побудови і розв'язання оптимізаційних задач в оцінці нерухомого майна.

У першому розділі розглядаються питання щодо організації оптимізаційного моделювання в оцінці нерухомого майна, видів моделей, основних етапів моделювання, випадкових подій і величин, їх числових характеристик, задач безумовної та умовної оптимізації та методи їх розв'язування.

Другий розділ присвячено практичним аспектам вирішення задач лінійного та нелінійного програмування в оцінці нерухомого майна, критеріям оцінки достовірності результатів дослідження, аналізу лінійних моделей оптимізаційних задач, інтерпретації отриманих результатів на основі аналізу лінійних моделей оптимізаційних задач.

У третьому розділі розглядаються поняття, сутність і зміст невизначеності та ризику під час оцінки нерухомого майна, досліджується питання щодо системи показників кількісного оцінювання ступеня ризику в оцінці нерухомого майна.

У четвертому розділі досліджуються питання економетричного моделювання в оцінці нерухомого майна, критерії адекватності економетричної моделі, сутність мультиколінеарності, напрями її виявлення, парна лінійна регресія та лінійні моделі множинної регресії.

У п'ятому розділі запропоновані і обґрунтовані методичні засоби нечіткого логічного виведення під час оперативного оцінювання станів складних об'єктів нерухомого майна за допомогою використання нечітких інтервальних уявлень про простір станів складних об'єктів; запропоновані і обґрунтовані методичні засоби настроювання параметрів функцій належності, які є результатом подальшого розвитку метода бінарного пошуку і покращують якість настроювання параметрів функцій належності.

РОЗДІЛ 1

ОРГАНІЗАЦІЯ ОПТИМІЗАЦІЙНОГО МОДЕЛЮВАННЯ В ОЦІНЦІ НЕРУХОМОГО МАЙНА

1.1 Концептуальні аспекти оптимізаційного моделювання в оцінці нерухомого майна

Математичні методи і моделі в оцінці нерухомості – це дисципліна, в якій формується система знань з методології та інструментарію побудови і використання різних типів оптимізаційних моделей в оцінці нерухомого майна. Теоретичним базисом математичних методів і моделей в оцінці нерухомості є системний аналіз, математичні закони, статистичні методи та інші.

Моделювання процесів в різних сферах діяльності людини здійснювалось ще з глибокої давнини [1]. Методи моделювання використовувались стародавніми єгиптянами і греками в будівництві й архітектурі, технічному конструюванні. Намагання оцінити процеси діяльності людини здійснив грецький філософ Аристотель. Проте формування методологічних аспектів і теоретичного базису використання математичних методів в оцінці нерухомого майна з'явилися значно пізніше.

Обґрунтування використання математичних методів починається з У. Петі (1623 – 1687 роки). В передмові до «Політичної арифметики» він вказує на те, що «замість того, щоб використовувати слова» необхідно перейти до «мови чисел, важелів і мір» [1]. Слід вказати, що моделювання процесів в суспільному виробництві здійснив Ф. Кене (1694 – 1774 роки) [2]. Чернишевський Н. Г. (1828 – 1889 роки) в замітках до трактату Д. Міля вказує на використання математичних методів [3].

На початку 19 століття стрімкий розвиток промислового виробництва призвів до необхідності використання математичних методів в різних сферах діяльності людини. З'являється «математична школа» засновником якої був О. Курно (1801 – 1877 роки). Представники математичної школи Г. Госсен (1810 – 1859 роки), В. Джевонс (1835 – 1882 роки), Л. Вальрас (1834 – 1910 роки), Г. Кассель (1866 – 1944 роки), Ф. Еджворт (1845 – 1926 роки), В. Парето (1848 – 1923 роки), В. Дмитрієв (1868 – 1913 роки) розробили методологію та сформували систему проведення різних досліджень математичними методами.

У 20-му столітті і до теперішнього часу математичні методи широко використовуються в різних сферах діяльності людини.

Ретроспективний аналіз розвитку математичного моделювання свідчить про його важливість і необхідність використання в різних процесах, адаптуючи його до сучасних трансформаційних умов. При моделюванні процесів оцінки нерухомості за допомогою математичних методів важливого значення набуває визначення моделі.

Термін модель походить від латинської означає «зразок, норма, міра». По суті модель представляє собою аналогію, подібність явищ, процесів, об'єктів.

Таким чином, **модель** – це зображення об'єкту або системи в оригінальній або аналоговій формі, якими управляють і використовують для прийняття управлінських рішень [4]. Прикладом моделей можуть бути макети машин, споруд, механізмів та інші. Ці моделі засновані на прямій подібності об'єктів.

В процесах оцінки нерухомості широко використовують моделі які побудовані на схожості поведінки системи, подібності їх реалізації на зміну дій.

Виділяють такі види моделей [5]: фізичні, символні, аналітичні, словесно-описові, математичні, імітаційні, формальні, функціональні, теоретичні.

Математичне моделювання здійснюють за такими етапами [5]:

- виявлення проблемних аспектів і особливостей досліджуваного процесу, його аналіз, ідентифікація і визначення достатньої структури для моделювання, формування мети і задач моделювання;
- аналіз досліджуваних процесів, оцінка використаних ресурсів і потужностей необхідних для здійснення процесу оцінки нерухомості;
- формування і обробка інформації щодо оцінки нерухомості;
- побудова моделі на основі математичного інструментарію;
- інтерпретація отриманої математичної моделі, уточнення і корегування її параметрів.

З позиції теорії пізнання спостережувані в природі й суспільстві *явища можна підрозділити на такі види* [6]:

- достовірні (визначені), які обов'язково відбудуться, якщо буде здійснена певна сукупність умов, і приймуть умови, які явно можна передбачити;
- неможливі, які явно не відбудуться в певних умовах;
- випадкові, які при сукупності умов можуть відбутися або не відбутися, в результаті випробувань можуть прийняти будь-яке значення, причому невідомо, яке саме;
- невизначені, про які нічого не можна сказати відбудуться вони або не відбудуться, незалежно від створених умов.

Слід розрізняти випадкові події-факти і випадкові величини.

Під «*подією*» розуміється будь-яке (не обов'язково знаменне) явище. *Подія-факт* в кількісному і якісному відношенні може бути величиною невизначеною, оскільки про неї нічого не можна сказати наперед відомою вірогідністю, а випадкова величина пов'язана з характером, змістом досліджуваного процесу [7].

Величина називається випадковою, якщо вона формується під дією багатьох дрібних причин, не піддатливих до результату випробувань повному контролю і обліку, діючих відносно незалежно один від одного.

Математичне моделювання вивчає кількісні закони масових випадкових величин і явищ, але не ставить перед собою завдання передбачити, відбудеться одинична подія чи ні.

Одним з напрямів математичного моделювання є вивчення закономірностей масових однорідних випадкових подій. Випадкові величини підрозділяються на *дискретні (переривчасті) й безперервні* [8].

Дискретними називаються випадкові величини, які приймають окремі, строго визначені, ізольовані, кінцеві чисельні значення з певною вірогідністю, між якою не може бути проміжних. При цьому число можливих значень дискретної випадкової величини може бути кінцевим і нескінченним.

Частіше зустрічаються *безперервні випадкові величини*, які можуть мати всі можливі значення в деякому кінцевому або нескінченному проміжку. Очевидно, що число можливих значень безперервної випадкової величини нескінченне.

Оскільки точність вимірювання або обліку завжди обмежена, то практично всі випадкові величини є дискретними.

Значна частина математичного моделювання пов'язана з необхідністю досліджувати і описувати велику сукупність об'єктів. Звичайно цю сукупність називають *генеральною*. Вона охоплює, наприклад, усіх мешканців великого міста, продукцію галузі народного господарства.

Якщо досліджувана сукупність об'єктів суттєво численна або об'єкти вивчення труднодоступні, а також є інші причини, що не дозволяють вивчити всі об'єкти, то вдаються до вивчення якоїсь частини генеральної сукупності, що називається *вибіркою* [9].

Вибірка повинна бути ваговою або, як кажуть, репрезентативною. Якщо вибірка представляє не всю генеральну сукупність, а якусь її частину, то це називається зсувом вибірки. Зсув – одне з основних джерел помилок при використанні вибіркового методу [9].

В математичному моделюванні необхідно визначити вибірку, що складається з n однорідних одиниць (елементів). Числом n називається обсяг вибірки. Одиницями вибірки можуть бути різні процеси і явища.

Чисельні значення, які приймає досліджувана ознака, називають варіантами. Зміна величини ознаки в статистичній сукупності називається *варіацією (вона коливається або розсіюється)* [9].

Для того, щоб замінити в зареєстрованих значеннях процесу, що вивчається, будь-яку закономірність, значення потрібно привести до доступного для аналізу вигляду, тобто впорядкувати, класифікувати, систематизувати, згрупувати. Процес розчленовування досліджуваної сукупності на частини називається *угрупованням* [9].

Початковою базою для вивчення закономірностей тих чи інших явищ служать статистичні ряди розподілу, які будуються за якісними і кількісними ознаками.

Якщо ряди є послідовністю чисел, що характеризують зміну показника в часі, то вони називаються тимчасовими, а якщо показують розподіл одиниць сукупності, що вивчається, по окремих групах, виділених за певною ознакою, то варіаційними.

Число, що показує, скільки разів зустрічається та або інша одиниці сукупності, що вивчається, називається його частотою m .

Варіаційний ряд або ряд розподілу є таблицею, в якій в порядку спадання або зростання перераховані можливі значення випадкової величини за вибіркою, що вивчається, з вказівкою їх частот (табл. 1.1).

Таблиця 1.1 – Значення випадкової величини

Значення випадкової величини	x_1	x_2	x_3	...	x_n
Частоти	m_1	m_2	m_3	...	m_n

Очевидно, що сума всіх частот дорівнює обсягу вибірки (1.1):

$$m_1 + m_2 + m_3 + \dots + m_n = \sum_{i=1}^n m_i = n, \quad (1.1)$$

де n – загальне число спостережень.

Якщо різних значень випадкової величини багато, то ряди розподілу складаються в інтервальній формі. Різниця між верхньою x_i і нижньою x_{i-1} межами інтервалу називається його величиною:

$$\Delta x_i = x_i - x_{i-1}. \quad (1.2)$$

Число інтервалів не повинно бути надмірно великим. Для того, щоб краще проявити характерні особливості, пов'язані з природою величин, що вивчаються, рекомендується ділити проміжок варіації на $6 \div 16$ інтервалів залежно від обсягу вибірки.

Отже, для визначення величини інтервалів необхідно різницю x_{\max} і x_{\min} (розмах варіювання) розділити на $6 \div 16$ залежно від числа спостережень:

$$\Delta x_i = \frac{x_{\max} - x_{\min}}{6 \div 16}. \quad (1.3)$$

У літературі зустрічається і таке визначення розрахунку інтервалів, як використання формули Стерджеса (1.4):

$$\Delta x = \frac{x_{\max} - x_{\min}}{1 + 3.32 \ln n}. \quad (1.4)$$

Отримане значення інтервалу Δx , звичайно, округляють. За величину і особливо центр інтервалу приймається деяке «зручне число», що має невелике число значущих цифр, щоб полегшити надалі обчислення.

Запис інтервальних рядів розподілу поданий в таблиці 1.2.

Таблиця 1.2 – Інтервальні ряди розподілу

Значення випадкових величин	$x_1 - x_2$	$x_2 - x_3$	$x_3 - x_4$...	$x_{n-1} - x_k$
частоти	m_1	m_2	m_3		m_k

Сума частот представлена таким чином:

$$\sum_{i=1}^k m_i = n. \quad (1.5)$$

Дослідження не обмежується побудовою ряду розподілу тієї або іншої випадкової величини. Необхідно знайти декілька величин, так званих статистичних характеристик, які відображали б властивості ряду розподілу в цілому, повніше характеризували б сукупність і властиві їй закономірності. Інакше кажучи, ставиться завдання знаходження такого значення випадкової величини, навколо якої групуються всі інші і зустрічаються найбільш часто.

Найважливішим і найпоширенішим з них є середня арифметична, яка позначається символом \bar{x} . Якщо випадкова величина приймає n значень, то середня арифметична за незгрупованими даними є сумою її значень, розділеною на їх число:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}. \quad (1.6)$$

Середня арифметична зважена за згрупованими даними обчислюється таким чином [10]:

$$\bar{x} = \frac{x_1 m_1 + x_2 m_2 + x_3 m_3 + \dots + x_k m_k}{\sum_{i=1}^k m_i} = \frac{\sum_{i=1}^k x_i m_i}{\sum_{i=1}^k m_i}. \quad (1.7)$$

Середня арифметична, як і середня арифметична зважена, володіє тією властивістю, що сума відхилень значень випадкової величини від середньої арифметичної дорівнює нулю. Ця властивість дає можливість визначити середню арифметичну як центр угруповання випадкової величини.

Крім середнього значення ознаки важливо знати характер варіації, тобто як тісно концентруються всі значення елементів сукупності навколо середньої. З теоретичної точки зору самою відповідною мірою коливання ознаки служить *дисперсія* (від латинського *dispercia* – розсіяння) [10], є квадратом відхилення

дослідних даних від середнього значення, за незгрупованими даними (1.8) та за згрупованими даними (1.9):

$$\sigma_x^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2}{n}, \quad (1.8)$$

$$\sigma_x^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 m_i}{n}. \quad (1.9)$$

Як бачимо, дисперсія випадкової величини в окремих випадках може мати нереальну розмірність. Для її усунення вводять середньоквадратичне відхилення, що розглядається в тих же одиницях вимірювання [10]:

$$\sigma_2 \sqrt{\sigma_x^2}. \quad (1.10)$$

Безрозмірним показником коливання випадкової величини є коефіцієнт варіації, який є відношенням σ_x до \bar{x} , відображеним у відсотках:

$$v_x = \frac{\sigma_x}{\bar{x}} * 100\%. \quad (1.11)$$

Якщо дисперсія або коефіцієнт варіації великі, то це свідчить про значний розкид середнього значення випадкової величини.

Випадкова величина характеризується двома основними параметрами:

- множиною її можливих значень;
- вірогідністю того, що вона прийме ті чи інші значення з цієї множини.

Вірогідність є одним з основних понять математичної статистики. Вона є математичним визначенням об'єктивної можливості відбутися або не відбутися випадковому явищу.

При вивченні рядів розподілу використовують не тільки абсолютні значення появи випадкової величини m_i , але й відносні частоти, тобто $\frac{m}{n}$.

Згідно з теоремою Якова Бернуллі (1654-1705 роки), що отримала назву «Закону великих чисел» в статистиці, можна передбачати відносну частоту події [11].

Теорема Я. Бернуллі була опублікована в 1713 році. Стосовно вибірки вона формулюється так: з вірогідністю, скільки завгодно близькою до одиниці, можна стверджувати, що різниця між відносною частотою і часткою в генеральній сукупності при достатньо великому обсязі вибірки буде скільки завгодно мала (1.12):

$$\frac{m}{n} = \frac{\text{вер}}{m \rightarrow \infty} p. \quad (1.12)$$

З цього виходить, що вірогідність події А визначається формулою:

$$P(A) = \frac{m}{n}. \quad (1.13)$$

Сума всіх відносних частот дорівнює 1, тобто

$$\frac{m_1}{n} + \frac{m_2}{n} + \dots + \frac{m_k}{n} = P_1 + P_2 + \dots + P_k = \sum_{i=1}^k P_i = 1. \quad (1.14)$$

З визначення вірогідності випливають такі її властивості [12]:

1. Вірогідність неможливої події дорівнює 1. При $m = n$:

$$P(A) = \frac{m}{n} = \frac{n}{n} = 1.$$

2. Вірогідність неможливої події дорівнює нулю. При $m = 0$:

$$P_a = \frac{m}{n} = \frac{0}{n} = 0.$$

3. Вірогідність випадкової події є позитивним числом, укладеним між нулем і одиницею.

Дійсно, випадковій події сприяє лише частина із загального числа елементарних результатів випробувань. У цьому випадку $0 < m < n$, значить:

$$0 < \frac{m}{n} < 1,$$

$$0 \leq P(A) \leq 1.$$

4. Вірогідність протилежної події дорівнює різниці між одиницею і вірогідністю даної події, тобто

$$P(B) = 1 - P(A).$$

5. Якщо в результаті випробувань повинно відбутися одне, і тільки одне з деяких подій A_1, A_2, \dots, A_k , то сума всієї вірогідності дорівнює одиниці, тобто

$$P_1(A_1) + P_2(A_2) + \dots + P_k(A_k) = 1.$$

Одним з узагальнюючих результатів закону великих чисел є те, що при достатньо великому числі спостережень n середнє значення випадкової величини \bar{x} приблизно дорівнює її математичному очікуванню, або $M_{(x)} = \bar{x}$. Така середня називається стохастичною [12].

З цього виходить, що математичне очікування дискретної випадкової величини є не випадкова (постійна) величина. П. Л. Чебишев (1821 – 1894 роки) довів, що сукупні дії великого числа чинників призводять до результату, майже не залежного від випадку [13].

У вузькому значенні під законом великих чисел розуміється ряд математичних теорем, в яких встановлюється факт наближення середніх показників у результаті великого числа спостережень до деяких постійних величин.

У широкому значенні зміст закону великих чисел полягає в тому, що при великому числі випадкових явищ їх середній результат практично перестає бути випадковим і може бути представлений з великою визначеністю.

Законом розподілу випадкової величини називається співвідношення, що встановлює зв'язок між можливими значеннями випадкової величини і відповідними їм ймовірностями.

Найпростішою формою завдання такого закону служить таблиця, в якій перераховані можливі значення випадкової величини і відповідні їм ймовірності (табл. 1.3).

Таблиця 1.3 – Значення випадкової величини і відповідні їм ймовірності

x_1	x_2	x_3	...	x_n	Разом
P_1	P_2	P_3	...	P_n	$\sum_{i=1}^n P_i = 1$

Щоб надати ряду розподілу наочний вигляд, будують його графічне зображення у вигляді гістограми, полігону, кумуляти і огіви [13].

Табличний розподіл можливих значень випадкової величини і відповідних їй ймовірностей, графічне зображення кривих розподілу і аналітичний опис вказаної залежності є формою закону розподілу [10 – 13].

Криві розподілу можуть бути різної форми. Проте серед них слід виділити так звані одновершинні криві, що часто зустрічаються. У дослідженнях симетричні розподіли зустрічаються рідко. Набагато частіше вершина кривої знаходиться не в центрі, а дещо зміщена.

Зустрічається також двопіковий розподіл. Його наявність свідчить про те, що розглядається неоднорідна сукупність.

Теоретичними розподілами у дослідженнях, головним чином, є закон Пуассона, показовий, біноміальний, Ст'юдента, χ^2 -квадрат, Лапласа, нормальний та інші [14].

Нормальний закон розподілу реалізується для випадкових величин, які формуються під сумарною дією багатьох незалежних поміж собою дрібних причин, дія кожної з яких мала в порівнянні із загальним результатом.

У математичній статистиці нормальний розподіл відіграє роль стандарту, з яким порівнюються інші розподіли.

Формула нормальної кривої має такий вигляд [13]:

$$Y = f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma_x^2}}, \quad (1.15)$$

де x – випадкова величина;

\bar{x} – середнє арифметичне або математичне очікування;

σ_x – середнє квадратичне відхилення;

$\pi=3,14159$, $e=2,71828$ – відомі константи.

Крива Гауса-Лапласа має горбоподібний вигляд і симетрично розташовується відносно вертикальної прямої (рис. 1.1) [13]. Центр угруповання випадкової величини і форму нормальної кривої визначають числові характеристики \bar{x} і σ_x .

При $x = \bar{x}$ функція має максимум, рівний (1.16):

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}}. \quad (1.16)$$

Симетрія кривої $x = \bar{x}$ вважається основною властивістю нормального розподілу: однакові відхилення значення випадкової величини від її середнього в обидві сторони зустрічаються однаково часто.

При збереженні своєї загальної форми крива розподілу нормального закону може мати різний ступінь пологості й крутизни залежно від значення σ_x (рис. 1.1).

У математико-статистичних дослідженнях, незалежно від розмірності випадкової величини x , може бути визначена відносна частота.

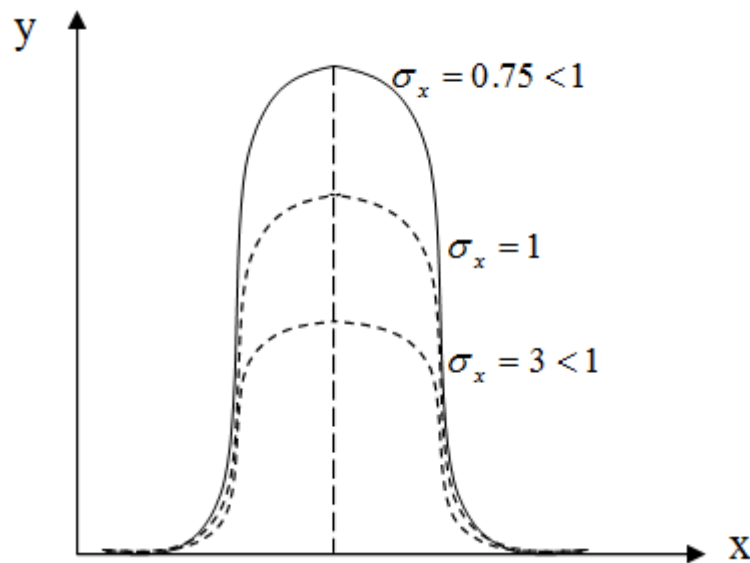


Рисунок 1.1 – Крива Гауса-Лапласа

По правому 3σ величина абсолютного відхилення випадкової величини від середнього по вибірці менше $\pm 3\sigma_x$ з вірогідністю 0,997. Лише 0,3 % всього x_i числа спостережень виходить з «трисигмових меж».

В інтервалі від $x - \sigma_x$ до $x + \sigma_x$ знаходиться 68,3% спостережень, в інтервалі від $x - 2\sigma_x$ до $x + 2\sigma_x$ – 95,5 % спостережень. Таким чином, максимум

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}}. \quad (1.17)$$

Оскільки площа диференційованої функції нормального розподілу дорівнює одиниці, то зі зростанням σ_x максимальна ордината нормальної кривої спадає, а сама крива стає більш пологою. Навпаки, зі спаданням σ_x нормальна крива стає більш гостроверхою.

При $\bar{x} = 0$ і $\sigma_x = 1$ нормальну криву називають нормованою [14]:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad (1.18)$$

Величина $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ табульована і може бути визначена з відповідних математико-статистичних таблиць (диференціальна функція Лапласа).

У цих таблицях наведені функції $f(x)$, відповідні позитивним значенням x . Для від'ємних x користуються тими ж таблицями, оскільки функція $f(x)$ парна, тобто $f(-x) = f(x)$. У таблиці наводяться значення $f(x)$ для x від 0 до 4 через 0,01.

Для того, щоб можна було користуватися готовими таблицями, потрібно криву нормального розподілу привести до стандартної форми. Стандартизація полягає в переході від випадкової величини x , що має математичне очікування \bar{x} і середньоквадратичне відхилення σ_x , до допоміжної величини, що називається центрованим і нормованим відхиленням [14]:

$$t = \frac{x - \bar{x}}{\sigma_x} \text{ чи } \Delta X = t \sigma_x. \quad (1.19)$$

Використовуючи відповідні таблиці значень, будують таблицю стандартизованого розподілу вірогідності.

Якщо на вісь абсцис нанести значення t , а на вісь ординат вірогідність $P(t)$, то графічне зображення дає нормальну криву. Фізичне значення t означає, на скільки середньоквадратичне відхилення σ_x змінює значення випадкової величини від її середнього значення \bar{x} .

При вибіркових обстеженнях допускаються різного роду похибки, при цьому розрізняють грубі, систематичні й випадкові помилки [15].

Грубі помилки за абсолютними величинами значно відрізняються від всього ряду помилок і підлягають виключенню з ряду спостережень.

Систематичні помилки є наслідком впливу певних чинників, що спотворюють результати вимірювань за певним законом у відомому напрямі. Вони викликані зносом засобів вимірювання, їх неправильною установкою, дією зовнішнього середовища і т.п.

Випадковими називають помилки, характер зміни яких не володіє видимою закономірністю. Кожна подальша помилка за абсолютним значенням може бути більше або менше попередньої.

Аналіз випадкових похибок ґрунтується на теорії випадкових помилок, яка дає змогу з певною гарантією вірогідністю обчислити дійсне значення шуканої величини.

В основі теорії випадкових помилок лежать такі підтверджені досвідом висновки [16]:

1. Різниця у значеннях характеристики вибіркової і генеральної сукупності складе помилку вибірки $\tilde{x} - \bar{x} = \varepsilon_x$, $\bar{\sigma} - \sigma_x = \varepsilon_\sigma$. Ці помилки є випадковими величинами.

Тому необхідно в кожному конкретному випадку визначити не тільки розмір помилки, але і надійність або гарантію того, що цей розмір не буде перевищений.

2. Вибіркові середні значення також симетрично розподіляються навкруги генеральної середньої, незалежно від характеру розподілу випадкової величини в генеральній сукупності.

Закономірність розподілу випадкових помилок спостережень описується нормальною кривою (рис. 1.2). Карл Гаус (1777 – 1855 роки) використовував її, як основу для теорії випадкових помилок вимірювань [14].

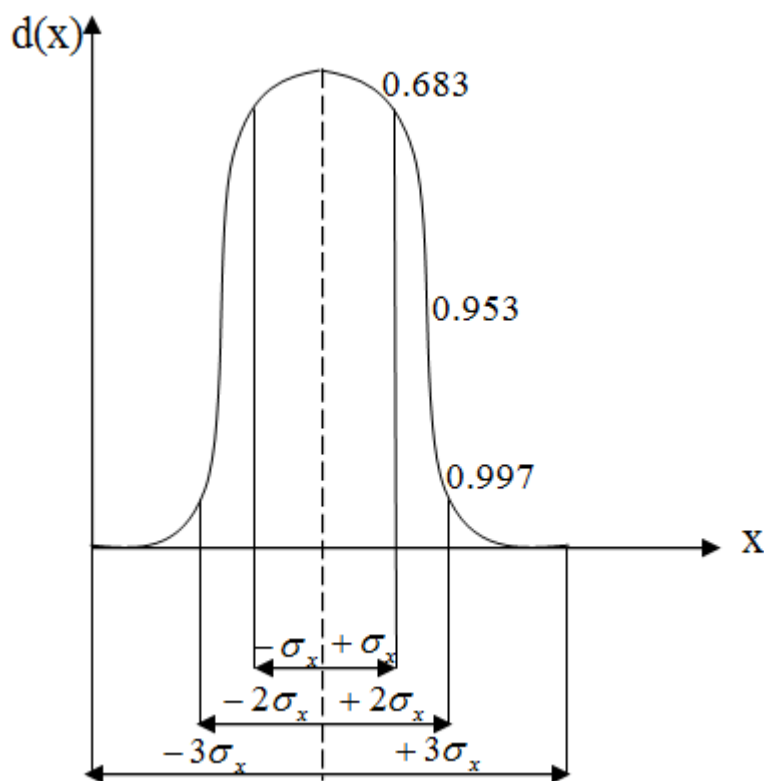


Рисунок 1.2 – Крива розподілу випадкових помилок спостережень

Вся площа під кривою дорівнює 1. Основна маса випадкових помилок групується навколо середнього значення, яке дорівнює 0.

На ділянці, обмеженій $\bar{x} + \sigma_x$, знаходиться 68,3 % всіх спостережень; на ділянці, обмеженій $\bar{x} + 2\sigma_x$ і $\bar{x} - 2\sigma_x$ – 95,3 %; на ділянці, обмеженій $\bar{x} + 3\sigma_x$ і $\bar{x} - 3\sigma_x$ – 99,7 %.

На основі характерних властивостей розподілу випадкових помилок спостережень можна зробити висновок, що при достатньо великому обсязі вибірки n її числові характеристики за вірогідністю наближаються до відповідних значень характеристики генеральної сукупності.

Рівень значущості вимірюється у відсотках і його чисельне значення засноване на так званому принципі незалежності маловірогідних подій. На практиці приймають рівні значущості, що знаходяться між 0,01-0,05. Відповідно їх називають одновідсотковими, двовідсотковими і т.п.

З принципу неможливості маловірогідних подій випливає такий висновок: якщо випадкова подія має вірогідність дуже близьку до одиниці, то можна вважати, що в одиничному випробуванні ця подія наступить ($P \geq 0,99$).

Припущення щодо закономірностей, які мають місце в генеральній сукупності, називається статистичною гіпотезою, а критерій її перевірки – статистичною характеристикою.

Критерієм перевірки вибирається деяка статистична характеристика. Припущення, що висувається, може бути помилковим, внаслідок вибіркової помилки, і має назву нульової гіпотези H_0 .

Конкуруюча гіпотеза означає, що має місце суттєва відмінність між вибірковими значеннями в генеральній сукупності. Сформулювавши гіпотезу H_0 , можна зіткнутися з чотирма ситуаціями:

- гіпотеза H_0 правильна, а її забракували, оскільки характеристика потрапила в критичну область, тобто допущена помилка першого роду, вірогідність якої рівна рівню значущості α ;
- гіпотеза правильна і її прийняли, оскільки характеристика потрапила в допустиму область, тобто рішення правильне;
- гіпотеза неправильна і її прийняли, оскільки характеристика потрапила в критичну область, тобто рішення правильне;
- гіпотеза неправильна, а її відкинули, оскільки характеристика потрапила в допустиму область. Допущена помилка другого роду, тобто прийнята невірна гіпотеза.

Як видно, рівень значущості можна тлумачити як ризик вчинити помилку першого роду, тобто забракувати правильну вірну гіпотезу. У зв'язку з цим для

ухвалення гіпотези рівень значущості призначають п'ятивідсотковий ($\alpha=0,05$), а для бракування гіпотези – одновідсотковий ($\alpha=0,01$).

Областю ухвалення гіпотези називають сукупність значень вибраного критерію, при яких гіпотезу приймають, а критичною областю – при значеннях критерію, коли нульову гіпотезу відкидають (рис. 1.3) [17].

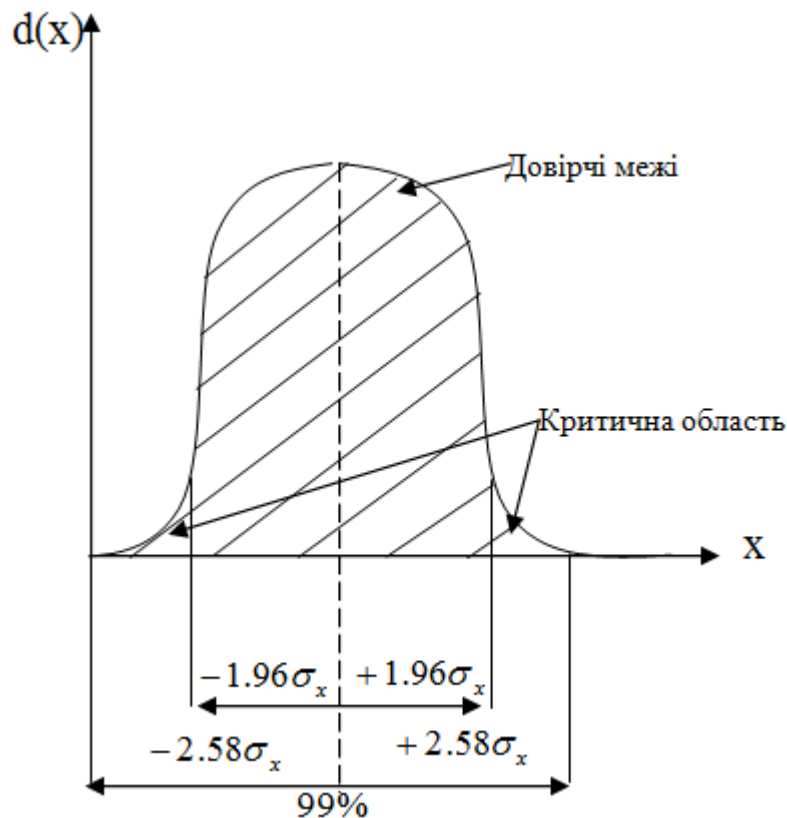


Рисунок 1.3 – Довірчі межі й критична область ряду розподілу

Вибіркове середнє є певне число, яке можна розглядати як випадкову величину. Отже, можна говорити про її розподіл і про числові характеристики цього розподілу (\bar{x} , σ_x^2 , σ_x та інші) [13].

Двостороння область, в яку повинне потрапити середнє значення генеральної сукупності, визначається довірчими межами при певному рівні значущості $2\Phi(t) = 0,95$ або за інтегральною функцією Лапласа $t = 1,96$.

Інтервал $x \pm 1,96\sqrt{\frac{\delta x^2}{n}}$ означає, що з вірогідністю $P = 0,95$ генеральна середня потрапляє в довірчі межі, а з вірогідністю $P = 0,05$ лежить зовні цих меж, тобто потрапляє в критичну область. Міра можливої відмінності між

вибірковою середньою і середньою генеральної сукупності має назву *стандартної помилки* [3].

Слід мати на увазі, що вибіркові середні, як і випадкові помилки спостережень, симетрично розподіляються навколо генеральної середньої за умови, що обсяг вибірки складає $n \geq 30$ спостережень. При малому обсязі вибіркових даних розподіл вибіркових середніх відрізняється від нормального розподілу тим сильніше, чим менше обсяг вибірки.

Межі довірчого інтервалу при малих вибірках обмежуються коефіцієнтом t_α , який був запропонований в 1908 році англійським математиком і хіміком В. С. Госсетом, що публікував свої роботи під псевдонімом «Ст'юдент» – студент. Надалі цей коефіцієнт отримав назву коефіцієнта Ст'юдента (це спеціально розроблені таблиці з урахуванням обсягу вибірки) [15].

Доцільно дотримуватися такої послідовності попередньої обробки результатів спостережень при $n \geq 30$ [9]:

- результати спостережень записують в таблицю;
- обчислюють середнє значення \bar{x} з n спостережень:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i; \quad (1.20)$$

- визначають похибки окремих спостережень:

$$\Delta x_i = \bar{x} - x_i \text{ та їх квадрати } (\Delta x_i)^2;$$

- відсіюють спостереження, що суттєво відмінні від інших, для цього:

- 1) знаходять середню квадратичну похибку:

$$\Delta \sigma_x = \sqrt{\frac{\sum_{i=1}^n \Delta x_i^2}{n}}; \quad (1.21)$$

- 2) задають значення похибки $\alpha = 0,95$;

- 3) визначають коефіцієнт Ст'юдента $t_\alpha = (n)$ для заданої надійності P і кількості спостережень n ;

- 4) знаходять межі довірчого інтервалу (похибки результатів спостережень):

$$\Delta x = t_\alpha(n) \Delta \sigma_x, x = \bar{x} \pm \Delta x;$$

5) обчислюють відносну похибку вибірових даних:

$$\varepsilon = \frac{\Delta x}{x} * 100\% . \quad (1.22)$$

Математичному моделюванню передус чітке уявлення суті задачі, що розв'язується, аналіз її змісту з використанням технологічної, економічної і інженерної логіки.

При цьому доцільно [5]:

- вивчити літературу і узагальнити професійні знання про об'єкт дослідження;
- чітко сформулювати мету і завдання дослідження;
- визначити джерела, обсяг і методи отримання інформації;
- провести попередній якісний і кількісний математико-статистичний аналіз результатів спостережень.

При визначенні умов, що впливають на досліджуваний показник, слід дотримуватися апробованих *принципів якісного аналізу* [7]:

- кожний чинник повинен бути теоретично обґрунтованим і змістовним, мати самостійне значення і не дублювати інші;
- вибірові дані повинні мати точне кількісне вимірювання, бути однорідними і співставними в часі й просторі.

Об'єктивність математичного моделювання багато в чому залежить від показовості (репрезентативності) й однорідності вибірових даних.

Заздалегідь обґрунтовується обсяг n -вибірки або перевіряється достатність початкової інформації для отримання математико-статистичних моделей заданої точності й надійності.

За теоремою Ляпунова для різних незалежних вибірок достатньо великого обсягу n , отриманих з однієї і тієї ж генеральної сукупності, середнє арифметичне \bar{Y} підпорядковується нормальному закону розподілу з дисперсією σ_y^2 , рівної $1/n$ -ї частини дисперсії випадкової величини. При цьому максимальне відхилення є вибірковою середньою \bar{Y} від генеральної середньої Y_i , що має назву стандартної помилки і визначається за формулою (1.23):

$$Y_i - \bar{Y} = t_\alpha \sqrt{\frac{\tau_y^2}{n}} , \quad (1.23)$$

де t_α – значення змінної в стандартизованому масштабі (1.24)

$$t_{\alpha} = \frac{Y_i - \bar{Y}}{\tau_y}. \quad (1.24)$$

Значення (1.24) визначається за інтегральною функцією Лапласа [3].
Таким чином,

$$n = \frac{t_{\alpha}^2 \cdot \tau_y^2}{\varepsilon^2}, \quad (1.25)$$

де n – кількість спостережень.

Розглянемо приклад. Необхідно встановити, при якому обсязі спостережень n -вибірка є генеральною сукупністю, якщо $P = 0,95$ або 95 %, $\varepsilon = 0,85$ і $\sigma_y = 4,56$?

Розв'язання

$P = 2\Phi(t_{\alpha}) = 0,95$ або $\Phi(t_{\alpha}) = \frac{0,95}{2} = 0,475$ за нормованою інтегральною функцією Лапласа знаходимо $t_{\alpha} = 1,96$.

Звідси,

$$n = \frac{t_{\alpha}^2 \cdot \tau_y^2}{\varepsilon^2} = \frac{1,96^2 \cdot 4,56^2}{0,85^2} = 35 \text{ спостережень.}$$

Виявлення спостережень, суттєво відмінних від основної маси вибірових даних, ґрунтується на тому, що коли \bar{Y}_i розподілені приблизно за нормальним законом, то найбільше відхилення від середнього значення за абсолютною величиною перевищує приблизно $3\sigma_y^2$, тобто всі спостереження повинні розміщуватися в інтервалі

$$\bar{Y} - 3\tau_y \leq y_i \leq \bar{Y} + 3\tau_y.$$

Точніше, контроль належності до досліджуваної вибірки суттєво відмінних значень проводиться при рівні значущості α з урахуванням обсягу вибірки n . При цьому визначається

$$\phi(t_{\alpha}) = \sqrt[n]{1 - \alpha} - 0,5,$$

а потім за таблицею інтегральної функції Лапласа знаходиться значення t_α , і допустимий інтервал записується у вигляді

$$\bar{Y} - t_\alpha \tau_y < Y_i < \bar{Y} + t_\alpha \tau_y.$$

Розглянемо приклад. Є вибірка обсягом $n = 150$ спостережень. Середнє значення по вибірці $\bar{Y} = 12,86$; середнє квадратичне відхилення $\sigma_y^2 = 6,24$; рівень значущості $\alpha = 0,05$; максимальне значення ознаки, що вивчається $y_{\max} = 32,64$; мінімальне – $y_{\min} = 3,42$. Визначити можливість використання в подальших дослідженнях y_{\max} і y_{\min} .

Розв'язання

При заданому рівні

$$\phi(t_\alpha) = \sqrt[150]{1 - 0,05} - 0,5 = 0,4996,$$

$$t_\alpha = 3,366$$

допустимий інтервал дорівнює

$$\begin{aligned} \bar{Y} - t_\alpha \tau_y = Y_i \leq \bar{Y} + t_\alpha \tau_y &= 12,86 - 3,366 \cdot 6,24 < Y_i < 12,86 + 3,366 \cdot 6,24 = \\ &= -8,14 < Y_i < 33,86. \end{aligned}$$

Всі спостереження можуть бути використані при подальшій обробці.

У разі, якщо початкова інформація отримана по декількох об'єктах або групах, необхідно перевірити її однорідність. Така перевірка ґрунтується на гіпотезі рівності вибірових середніх обсягами n_i і n_j , отриманих з однієї генеральної сукупності.

За теоремою Чебишева, зі збільшенням обсягу вибірки її середнє значення прямує за вірогідністю до генеральної середньої, тому впливає такий висновок: якщо по декількох вибірках достатньо великого обсягу з однієї і тієї ж генеральної сукупності буде знайдено вибіркові середні \bar{Y}_i і \bar{Y}_j , то вони будуть приблизно рівні між собою.

За умови незалежності вибірок і їх належності до єдиної нормально розподіленої генеральної сукупності для будь-яких двох вибірок i -ої і j -ої маємо ймовірність

$$\{|\bar{Y}_i - \bar{Y}_j|\} \leq t_{ij} \sqrt{\frac{\tau_i^2}{n_i} + \frac{\tau_j^2}{n_j}} = 2\phi(t_{ij}), \quad (1.26)$$

де σ_i^2, σ_j^2 – вибіркові дисперсії;

n_i, n_j – обсяги вибірок.

Різниці $\bar{Y}_i - \bar{Y}_j$ відносяться до відповідної стандартної помилки. Як критерій перевірки приймають нормовану різницю, яку обчислюють на основі співвідношення:

$$\frac{\bar{Y}_i - \bar{Y}_j}{\sqrt{\frac{\tau_i^2}{n_i} + \frac{\tau_j^2}{n_j}}} \leq t_{ij},$$

що порівнюється з табличним значенням t_α , де $2\Phi(t_\alpha) = 1 - \alpha$.

Гіпотеза однорідності вибірових даних підтверджується при $P = 2\Phi(t_\alpha) = 0,95$ і менше, тобто $\alpha = 0,05$ і більше. Це означає, що при всіх значеннях $t_{ij} \leq 1,96$ вся сукупність вихідних даних вважається приблизно однорідною і обробка може вестися по всьому масиву.

Розглянемо приклад. По двох об'єктах зібрана інформація з такими кількісними характеристиками: $n_1=54$; $n_2=56$; $\bar{Y}_1=16,13$; $\bar{Y}_2=13,5$; $\sigma_{y1}^2=65,3$; $\sigma_{y2}^2=57,9$. Необхідно визначити рівень значимості при формуванні гіпотези про однорідність сукупності вибірових даних.

Розв'язання

Визначаємо $t_{ij}(\max)$ для y_1 і y_2 :

$$t_{ij} = \frac{16,13 - 13,5}{\sqrt{\frac{65,3}{54} + \frac{57,9}{56}}} = 1,76.$$

Звідси, $P = 2\Phi(1,76) = 0,92$ або 92 %.

Гіпотеза про однорідність сукупності вибірових даних підтверджується з рівнем значущості $\alpha = 0,08$ або 8 %.

Необхідність знання закону розподілу в кореляційному аналізі зумовлена насамперед обґрунтуванням форми зв'язку між змінними.

Нормальний закон реалізується для випадкових величин, які формуються під сумарною дією багатьох відносно незалежних між собою причин, дія кожної з яких незначна в порівнянні із загальним результатом.

Результати спостережень обробляють в такій послідовності [10].

1. Вихідні дані розбиваються на інтервали і складають ряд розподілу функціональної ознаки y_i , визначають абсолютні й відносні частоти і будують гістограма розподілу.

2. Розраховують параметри закону розподілу \bar{Y} і σ_y . Для спрощення рахункової роботи вводиться безрозмірна величина

$$\bar{Y}' = \frac{Y_{i_{cp}} - C_y}{\Delta y}, \quad (1.27)$$

де $Y_{i_{cp}}$ – деяке інтервальне значення функції;

C_y – інтервальне значення $Y_{i_{cp}}$, прийняте за центр угруповання;

Δy – інтервал зміни випадкової величини.

Дійсне значення \bar{Y} і σ_y обчислюють на основі співвідношень $\bar{Y} = \bar{Y}' \Delta y + C_y$, $\tau_y^2 = \tau_y'^2 \cdot \Delta_y^2$ та $\tau_y = \tau_y' \cdot \Delta_y$.

3. Знаходять середнє інтервальне значення $Y_{i_{cp}}$ в стандартизованому масштабі, відповідне центрам інтервалів. За допомогою диференціальної функції Лапласа для кожного t_i знаходять значення $f(t)$.

4. Визначають ординати теоретичної кривої розподілу і за знайденими точками будують теоретичну криву:

$$Y_n = \frac{f(t)}{\tau_y}. \quad (1.28)$$

5. Оцінюють ступінь подібності теоретичної кривої з дослідженими даними. Оцінку ступеня згоди частіш за все проводять за допомогою

критерію χ^2 – «хі-квадрат» Пірсона, який є спеціально підібраною випадковою величиною, що визначається за формулою

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - \bar{m})^2}{\bar{m}_i}, \quad (1.29)$$

де k – число інтервалів угруповання змінної;

\bar{m}_i – емпіричні й теоретичні частоти.

6. Задаючись довірчим рівнянням значущості $\alpha = 5\%$, за допомогою таблиці χ^2 – розподілу за числом ступеней свободи

$$f = K - (S + 1), \quad (1.30)$$

де K – число інтервалів;

S – ступінь свободи (для нормального розподілу $S = 2(\bar{Y}, \sigma_y)$, оскільки необхідно скласти 2 рівняння для знаходження теоретичного розподілу \bar{Y} і σ_y).

7. Встановлюють критичне значення χ^2 , з якими порівнюють розрахункове значення.

Якщо обчислене значення χ^2 за дослідженими даними менше табличного, тобто воно потрапляє в область прийняття гіпотези H_0 , то теоретична крива розподілу узгоджується з емпіричним розподілом. Якщо чисельне значення χ^2 перевершує табличне або рівне йому, тобто воно потрапляє в критичну область, дана гіпотеза H_0 про форму кривої розподілу відкидається.

Розглянемо приклад. Необхідно визначити закон розподілу витрат часу проходження рухомим складом маршруту між двома зупинками (хв.) при $n = 180$ спостережень і $y_{\min} = 0,70$, $y_{\max} = 1,57$ хв. Розмір інтервалу складає 0,1. Необхідно побудувати гістограму і полігон розподілу, а також розрахувати показники нормального закону розподілу.

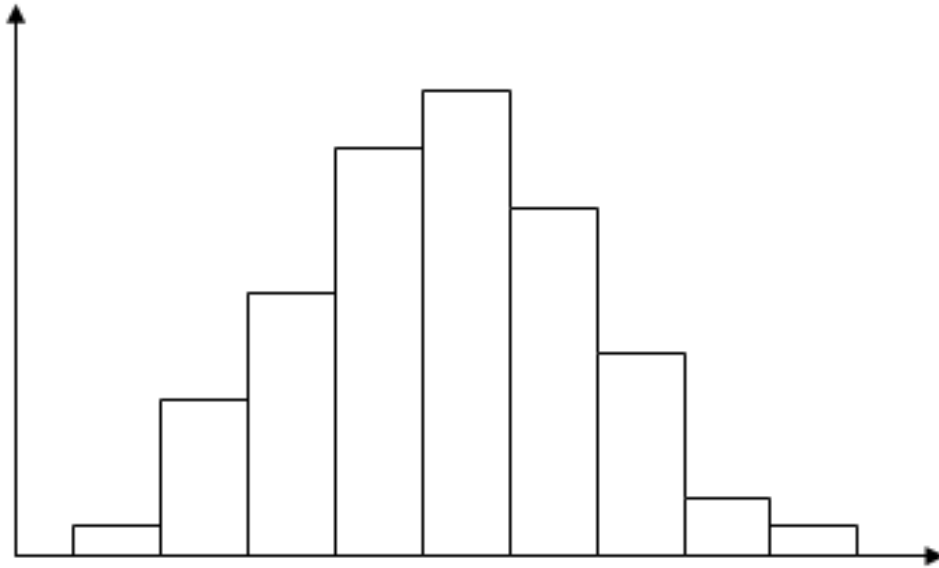


Рисунок 1.4 – Гістограма розподілу

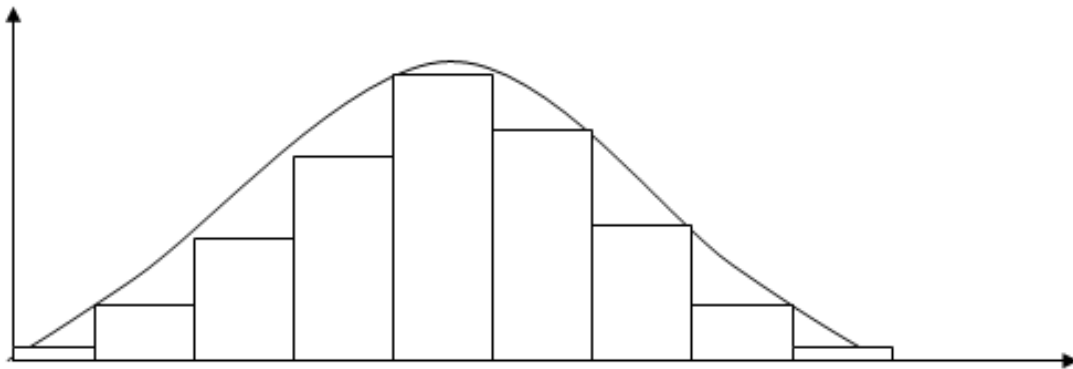


Рисунок 1.5 – Гістограма і полігон розподілу

На підставі даних, представлених у таблиці 1.4, отримуємо:

$$\bar{Y}' = \frac{\sum_{i=1}^k m_i y'_{cp}}{n} = -\frac{69}{180} = -0,38;$$

$$\bar{Y} = \bar{Y}' \Delta y + C_y = (-0,38) \cdot 0,1 + 1,15 = 1,11;$$

$$\bar{Y}_{cp}'^2 = \frac{\sum_{i=1}^k y_i'^2 m_i}{n} = \frac{481}{180} = 2,67;$$

$$\tau_y'^2 = \tau_{cp}'^2 - \bar{y}'^2 = 2,97 - (-0,38)^2 = 2,57;$$

$$\tau_y^2 = \Delta y^2 \tau_y'^2 = 0,1^2 \cdot 2,57 = 0,025;$$

$$\tau_y = \sqrt{0,025} = 0,159;$$

$$\tau_y'^2 = \frac{(y_{cp}' - \bar{y}')^2 m_i}{n} = \frac{448,6316}{180} = 2,4924;$$

$$\tau_y^2 = 0,1^2 \cdot 2,4924 = 0,025;$$

$$\tau_y = \sqrt{0,025} = 0,159;$$

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - \tilde{m}_i)^2}{m_i} = \frac{(3 - 3,36)^2}{3,36} + \frac{(13 - 11,9)^2}{11,9} + \dots + \frac{(1 - 1,03)^2}{1,03} = 0,795;$$

$$f = k - (S + 1) = 9 - (2 + 1) = 6;$$

$$\chi_{роз}^2 < \chi_{табл}^2 \text{ 5\%}.$$

Таким чином, теоретична крива розподілу зіставляється з емпіричним розподілом, що свідчить про наявність нормального розподілу.

Таблиця 1.4 – Розрахунок показників нормального закону розподілу

Інтервал Δy	Середнє значення інтервалу	Частота	Відносна частота	Умовні варіанти	Розрахунок середнього значення	Розрахунок дисперсії		Значення в стандартизованому масштабі	Значення диференціальної функції	Емпіричні розрахунки	Ординати теоретичного розподілу	Розрахункові частоти
$y_{k-1}-y_k$	y_{cp}	m_i	m_i/n	Y'_{cp}	$m_i y'_{cp}$	$m_i y'^2_{cp}$	$(y'_{cp} - \bar{y}'_{cp})^2 m_i$	$t_i = \left \frac{y_{cp} - \bar{y}}{\tau_y} \right $	$f(t)$	$y_2 = \frac{m_i}{n \Delta y}$	$y_n = \frac{f(t)}{\tau_y}$	$m_i = y_n \cdot n \Delta y$
0,7 – 0,8	0,75	3	0,017	-4	-12	48	39,3132	2,26	0,031	0,17	0,195	3,36
0,8 – 0,9	0,85	13	0,072	-3	-39	117	89,2372	1,63	0,1057	0,72	0,665	11,90
0,9 – 1,0	0,95	30	0,167	-2	-60	120	78,7320	1,00	0,2420	1,67	1,522	27,40
1,0 – 1,1	1,05	40	0,222	-1	-40	40	15,3760	0,38	0,3712	2,22	2,385	42,03
1,1 – 1,2	1,15	41	0,228	0	0	0	0	0,25	0,3867	2,28	2,432	43,7
1,2 – 1,3	1,25	31	0,172	1	31	31	59,0364	0,88	0,2709	1,72	1,704	30,7
1,3 – 1,4	1,35	16	0,089	2	32	64	90,6304	1,50	0,1295	0,89	0,817	14,7
1,4 – 1,5	1,45	5	0,028	3	15	45	57,1220	2,13	0,0413	0,28	0,259	4,66
1,5 – 1,6	1,55	1	0,005	4	4	16	19,1844	2,75	0,0091	0,05	0,057	1,03

1.2 Оптимізаційні моделі в оцінці нерухомого майна: напрями формування та особливості застосування

На якому б рівні не знаходилося суспільне виробництво, які великі не були б трудові, матеріальні й фінансові ресурси, перед господарськими керівниками завжди стоїть завдання найкращого використання виробничих ресурсів і потужностей. Тобто необхідно знайти оптимальне значення між представленими виробничими ресурсами. Для цього необхідно сформувати математичні моделі оптимізації шляхом розв'язання задачі оптимізації.

Розв'язати оптимізаційну задачу – це знайти оптимальне її розв'язання або встановити, що розв'язання немає. Методами розв'язання оптимізаційних задач є методи *математичного програмування* [9].

Уперше подібна задача у вигляді пропозиції щодо укладання національного плану перевезень, що дозволяє мінімізувати сумарний кілометраж, подана в роботі Л. М. Толстого, екстремальна задача з мінімізації транспортних витрат була ним сформульована в 1939 році [11].

Одну з різновидів транспортної задачі в 1941 році представив американець Хічкок (проблема Хічкока) [3]. Але закінченого методу вирішення цієї задачі він не розробив.

У загальному вигляді задача математичного програмування сформульована в 1939 році Л. В. Канторовичем. Він же запропонував метод множників, що дозволяє її вирішувати. Разом із М. К. Гавуриним у 1949 році Л. В. Канторович розробив метод потенціалів, який і дотепер є найбільш поширеним методом вирішення транспортних задач [3].

Широко відомий метод вирішення задачі лінійного програмування – симплексний метод – був опублікований Д. Б. Данцигом у 1949 році [4]. Вдалою модифікацією симплексного методу є диференціальний алгоритм, що логічно впливає з диференціального алгоритму вирішення загальної задачі математичного програмування.

Застосування математичних методів в оцінці землі та нерухомого майна на першому етапі ознаменувалося досить гострою дискусією фахівців «традиційної» школи та нового покоління. Однак, тепер мало залишилося фахівців, які б прямо заперечували проти необхідності використання ефективних математичних методів при вирішенні таких важливих проблем.

Основними економічними передумовами постановки і вирішення задач методами математичного програмування для формування математичних

моделей оптимізації слід вважати [8]:

- широке використання математичних методів у сполученні із сучасними засобами електронно-обчислювальної техніки;
- можливість одержання необхідної і достовірної інформації;
- достатньо повна теоретична розробка методів вирішення задач математичного програмування [4].

Математичне програмування відіграє винятково важливу роль у підготовці фахівців з оцінки землі та нерухомого майна. Використання математичних методів дозволяє вирішувати оптимальним способом багато задач організації, планування і управління. Іншими словами, це надійний інструмент для одержання найвищого ефекту в конкретних умовах.

Вираз «математичне програмування» слід розуміти, як ітераційний пошук найкращого варіанта використання обмежених потужностей і ресурсів для досягнення поставлених цілей [4].

У математиці максимум і мінімум мають назву – екстремум, а задачі пошуку екстремуму називають екстремальними задачами. Ті припустимі рішення, при яких досягається оптимум, називають оптимальними, або екстремальними рішеннями. У загальному випадку екстремальна задача може мати одне, декілька, множину, нескінченну множину або жодного оптимального рішення.

Змістовна постановка задачі повинна дозволяти переходити до строгої математичної моделі.

У загальному вигляді екстремальна задача формулюється таким чином: знайти найбільше (максимальне) або найменше (мінімальне) значення деякої функції

$$Y(x_1, x_2, \dots, x_n)$$

при умовах

$$f_i(x_1, x_2, \dots, x_n) \leq b_i (i = 1, m),$$

де Y і f_i – задані функції;

b_i – дійсне число.

Наведене формулювання є узагальненням постановок ряду часткових задач математичного програмування, що можуть розрізнятися між собою як видом функцій Y і f_i (лінійні, нелінійні, стохастичні), так і характером (дискретні, неперервні) змінних [14].

Функцію $Y(x_1, x_2, \dots, x_n)$, яку мінімізують або максимізують, називають цільовою функцією. Залежно від особливостей функцій Y і f_i математичне програмування можна розподілити на ряд самостійних дисциплін, що вивчають і розробляють методи вирішення окремих класів екстремальних задач.

Задачі математичного програмування розподіляються на задачі лінійного і нелінійного програмування. При цьому, якщо усі функції Y і f_i є лінійними, то відповідна задача відноситься до класу задач лінійного програмування. Якщо ж хоча б одна із зазначених функцій є нелінійною, то відповідна задача належить до класу задач нелінійного програмування.

Окремими класами задач математичного програмування є *задачі цілочислового, параметричного і дрібно-лінійного програмування* [16].

У задачах цілочислового, або дискретного програмування частина або всі невідомі можуть приймати тільки цілочислові значення.

У задачах параметричного програмування цільова функція або функції обмежень, що визначають область можливих змін змінних, залежать від деяких параметрів.

У задачах дрібно-лінійного програмування цільова функція є відношенням двох лінійних функцій, а функції, що визначають область припустимих рішень, також є лінійними.

Особливі класи становлять задачі стохастичного і динамічного програмування.

Стохастичне програмування використовують для вирішення задач, в яких обмеження мають імовірний чи випадковий характер, тобто, необхідно враховувати вплив яких-небудь непередбачених обставин. У задачах стохастичного програмування може бути математичне очікування деякого показника.

За допомогою лінійного, нелінійного, цілочислового і стохастичного програмування вирішуються задачі, що зводяться до відшукування оптимального рішення без урахування можливої динаміки процесу, тобто без урахування чинника часу.

У динамічному програмуванні мають місце багатоетапні задачі, що вимагають оптимізації прийнятих рішень не як одиничного акту, а з урахуванням розвитку явища, його зміни в часі.

Переваги динамічного програмування [6]:

- можливість поетапного аналізу результатів у процесі вирішення задачі, визначення оптимальної стратегії з урахуванням чинника часу;

- поглиблення раніше розроблених методів кількісного і якісного характеру дослідження процесів;
- більш об'єктивне, повне і точне вирішення завдань.

Таким чином, математичне програмування є важливим інструментарієм побудови математичних моделей оптимізації, що досліджує екстремальні задачі і розробляє методи вирішення. Математичне програмування як наука знаходиться в процесі постійного розвитку. Вченими всього світу розроблено багато методів для вирішення різних класів задач математичного програмування. Разом з тим багато задач ще не мають ефективних методів вирішення і чекають своїх дослідників.

Вирішення екстремальної задачі з оцінки землі та нерухомого майна складається з *таких етапів* [16]:

- побудови математичної моделі, тобто обґрунтування критерію оптимізації, виявлення і формалізації у вигляді системи рівнянь або нерівностей найбільш істотних обмежень задачі;
- вибору математичного методу, що дозволяє за кінцеве число кроків одержати шукане рішення з будь-якою заздалегідь заданою точністю, або вибору відповідної комп'ютерної технології;
- аналіз отриманих результатів з позицій можливого їхнього практичного застосування.

Критерієм оптимальності називається показник, за яким оцінюється міра ефективності плану. Критерій оптимальності повинен бути однозначним і мати кількісний вираз.

Для побудови математичної моделі найбільш часто використовуються задачі лінійного програмування. Тому розглянемо їх більш докладно.

Загальна задача лінійного програмування формулюється таким чином: знайти оптимум лінійної функції $y(x)$, якщо на змінні задачі накладені лінійні обмеження у вигляді рівностей і нерівностей.

Аналітичний запис цього завдання має такий вигляд [17]:

$$y(x) = c^T x + c_0 \rightarrow \text{opt}; \quad (1.31)$$

$$x \in \Omega \subset R^n;$$

$$\Omega: A_1 x + b_1 \leq 0; \quad (1.32)$$

$$A_2 x + b_2 = 0; \quad (1.33)$$

$$A_3 x + b_3 \geq 0; \quad (1.34)$$

$$x \geq 0, \quad (1.35)$$

де x – n -мірний вектор дійсних змінних;

c – n -мірний вектор коефіцієнтів;

c_0 – вільний член у складі функції y ;

A_1, A_2, A_3 – матриці коефіцієнтів лінійних систем розмірності $m_1 \times n$, $m_2 \times n$, $m_3 \times n$ відповідно, $m_2 < n$;

b_1, b_2, b_3 – вектори вільних членів обмежень розмірності $m_1 \times 1$, $m_2 \times 1$, $m_3 \times 1$ відповідно.

Часткові задачі лінійного програмування можуть не містити однієї або двох систем обмежень типу (1.32) – (1.34). Крім того, замість умови невід’ємності (1.35) може мати місце двостороння або одностороння обмеженість змінних.

Задачу, складену з (1.31), (1.32) і (1.35), називають стандартною задачею лінійного програмування.

Канонічна, або основна задача лінійного програмування має такий вигляд:

$$y(x) = c^T x + c_0 \rightarrow \max; \quad (1.36)$$

$$x \in \Omega \subset \mathbb{R}^n;$$

$$\Omega: A x + b = 0; \quad (1.37)$$

$$x \geq 0, \quad (1.38)$$

де A – матриця коефіцієнтів розмірності $m \times n$, $m < n$;

b – вектори вільних членів розмірності $m \times 1$.

Очевидно, що обмеження-нерівність типу « \leq » можна перетворити в обмеження-рівність додаванням до його лівої частини додаткової невід’ємної змінної, а кожне обмеження-нерівність типу « \geq » – в обмеження-рівність шляхом вирахуванням з його лівої частини додаткової невід’ємної змінної. Задачу мінімізації лінійної функції y можна звести до задачі максимізації шляхом множення останньої на -1 . Таким чином, задачу лінійної оптимізації (1.31) – (1.35) завжди можна перетворити в задачу (1.36) – (1.38) і навпаки.

Складання математичної моделі загальної задачі математичного програмування або її канонічної форми вимагає певних зусиль і кмітливості, досвід створення математичних моделей швидко накопичується. Досить мати практику вирішення декількох задач, щоб надалі не мати особливих труднощів при переході від змістовної постановки задачі лінійного програмування до формальної (аналітичної) [4].

У моделюванні використовують і розв'язують *задачі безумовної оптимізації*, в яких задається лише одна цільова функція. В задачах безумовної оптимізації не існує обмежень і граничних умов. У цих задачах поняття оптимуму та екстремуму збігаються, і для знаходження оптимуму в них застосовуються методи знаходження екстремуму. Слід зазначити, що найбільше або найменше значення – це екстремум, а оптимум – оптимальне найбільше або найменше значення.

У цих задачах знаходять першу похідну функції, прирівнюють її до нуля, знаходять параметри моделі, знаходять другу похідну і визначають її знак. Якщо друга похідна більша за 0, то точка x – мінімум функції.

Методами розв'язання задач умовної оптимізації є [6]:

- метод штрафних функцій, в якому мінімізується нова цільова функція, яка містить у собі першу цільову функцію та задані обмеження. При цьому визначається штрафна функція;
- метод Лагранжа полягає у побудові функції виду:

$$L(x_1, x_2, X) = f(x_1, x_2) + X g(x_1, x_2),$$

тобто, зведення задачі на умовний екстремум двох незалежних змінних до задачі на абсолютний екстремум функції

$$L\{x_1, x_2, X\}$$

трьох незалежних змінних x_1, x_2, X . Функція Лагранжа є сумою цільової функції та функції обмеження, помноженої на нову незалежну змінну X (множник Лагранжа), яка має перший порядок. Для знаходження точок умовного локального екстремуму функції за наявності обмеження слід насамперед знайти критичні точки функції Лагранжа. Критичні точки функції Лагранжа потрібно скоротити на координати X . Кожну одержану скорочену точку необхідно проаналізувати та з'ясувати чи є вона точкою умовного екстремуму функції за даних обмежень чи ні.

Питання та завдання для самоконтролю до розділу 1

Питання для самоконтролю:

1. Що таке математичне моделювання?
2. Назвіть етапи розвитку математичного моделювання.
3. Визначте поняття «модель», які види моделей Ви можете назвати.
4. Назвіть основні етапи моделювання.
5. Які види явищ Ви знаєте?
6. Визначте випадкову величину і її числову характеристику.
7. Назвіть і охарактеризуйте закони розподілу випадкової величини.
8. Як перевіряють статистичні гіпотези?
9. Назвіть етапи попередньої обробки інформації.
10. Охарактеризуйте оптимізаційні моделі і назвіть їх види.
11. У чому полягають задачі умовної і безумовної оптимізації?
12. Які методи використовуються для вирішення задач умовної і безумовної оптимізації, в чому вони полягають?

Завдання для самоконтролю:

1. Встановити, при якому обсязі спостережень n -вибірка є генеральною сукупністю, якщо $P = 0,95$ або 95% , $\varepsilon = 0,80$ і $\sigma_y = 4,18$?
2. Є вибірка обсягом $n = 100$ спостережень. Середнє значення по вибірці $\bar{Y} = 10,12$; середнє квадратичне відхилення $\sigma_y^2 = 5,12$; рівень значущості $\alpha = 0,05$; максимальне значення ознаки $y_{\max} = 26,16$; мінімальне – $y_{\min} = 3,09$. Визначити можливість використання в подальших дослідженнях y_{\max} і y_{\min} .
3. По двох об'єктах зібрана інформація з такими кількісними характеристиками: $n_1 = 45$; $n_2 = 46$; $\bar{Y}_1 = 15,17$; $\bar{Y}_2 = 12,5$; $\sigma_{y1}^2 = 61,4$; $\sigma_{y2}^2 = 55,6$. Визначити рівень значущості при формуванні гіпотези про однорідність сукупності вибірових даних.
4. Визначити закон розподілу витрат часу проходження рухомим складом маршруту між двома зупинками (хв.) при $n = 190$ спостережень і $y_{\min} = 0,50$ хв., $y_{\max} = 1,46$ хв. Розмір інтервалу складає $0,1$. Побудуйте гістограму і полігон розподілу. Розрахуйте показники нормального закону розподілу.
5. Знайдіть екстремум функції $y = x_1 + x_2$ за умови $x_1 + x_2 - 1 = 0$ або розв'яжіть задачу на умовний екстремум методом Лагранжа.

РОЗДІЛ 2

ЛІНІЙНЕ ТА НЕЛІНІЙНЕ ПРОГРАМУВАННЯ В ОЦІНЦІ НЕРУХОМОГО МАЙНА

2.1 Задача лінійного програмування та методи її розв'язання

Лінійне програмування використовує математичний інструментарій, який базується на теорії і методах вирішення задач про екстремуми лінійних функцій, що задаються системами лінійних рівнянь.

Найбільш універсальним методом лінійного програмування є *диференціальний алгоритм*, що логічно випливає з диференціального алгоритму загальної задачі математичного програмування. Диференціальний алгоритм, як і широко відомий симплекс-метод, дозволяє вирішувати будь-які задачі лінійного програмування. Однак, для деяких класів задач лінійного програмування доцільно використовувати більш прості методи. Так, для вирішення задач із кількістю змінних, рівною двом, використовують *графічний метод*, що відзначається простотою і наочністю, але потребує графічних побудов. Для вирішення задач лінійного програмування, відомих як транспортні, використовують *метод потенціалів* [18].

Методичною основою обчислювальних процедур будь-якого методу є принцип аналізу і послідовного поліпшення деякого початкового плану розподілу і використання ресурсів. План поліпшують доти, поки не буде знайдений найкращий (оптимальний) варіант. Іншими словами, спочатку складається деякий початковий план, що аналізується за конкретними строго розробленими правилами. На підставі аналізу визначаються можливість і напрямок поліпшення початкового варіанта плану. Потім обчислюється новий план, що піддається такому ж аналізу і подальшому поліпшенню, тобто наближенню до оптимуму. Обчислювальний процес продовжується доти, поки аналіз не покаже неможливість дальшого поліпшення.

Симплекс-метод використовується для вирішення будь-якої задачі лінійного програмування [19]. *Сутність симплекс-методу* полягає в тому, що, відправляючись з деякої довільної вершини багатокутника обмежень, переходять до обчислення тільки такої вершини, в якій значення лінійної форми буде більшим, ніж в попередній [19]. Решта варіантів не обчислюється. Так, при кінцевому порівняно малому числі кроків може бути знайдений оптимальний план. Таким чином, проводиться впорядкований перебір вершин, при якому відбувається постійне збільшення лінійної форми. В цьому аспекті

симплексний метод називається також методом послідовного поліпшення плану.

Вирішення задач лінійного програмування [19] симплекс-методом полягає:

- в розробці базового рішення на оптимальність. Якщо воно оптимальне, то задача вирішена, в іншому випадку виконують другий етап;

- визначаються вектор \vec{A}_k , який повинен бути введений в базис, і вектор \vec{A}_r , який повинен бути виключений з нього, тобто виходить новий базисний план з великим значенням лінійної форми. Щоб знайти вектори \vec{A}_k і \vec{A}_r , заміна яких забезпечує найбільше зростання лінійної форми, виразимо всі вектори, що не входять в базис, через базисні вектори

$$\vec{A}_j = \sum_{i=1}^m a_{ij} \vec{A}_i, \quad (2.1)$$

де a_{ij} – проекції вектора \vec{A}_j на вектор \vec{A}_i .

Запишемо систему обмежень у векторній формі в такому вигляді:

$$\sum_{i=1}^m x_i \vec{A}_i - \theta \vec{A}_k + \theta \vec{A}_k = \vec{B}. \quad (2.2)$$

Оскільки

$$\vec{A}_k = \sum_{i=1}^m a_{ik} \vec{A}_i,$$

то

$$\sum_{i=1}^m (x_i - \theta a_{ik}) \vec{A}_i + \theta \vec{A}_k = \vec{B}. \quad (2.3)$$

Співвідношення (2.2) дає рішення тільки у тому випадку, коли коефіцієнти при векторах \vec{A}_i і \vec{A}_k нового базису будуть додатними, тобто:

$$x'_i = x_i - \theta a_{ik} \geq 0,$$

$$\theta \geq 0. \quad (2.4)$$

Відповідне нове значення лінійної форми прийме вигляд:

$$L_1 = \sum_{i=1}^m (x_i - \theta a_{ik}) c_i + \theta c_k = \sum_{i=1}^m x_i c_i + \theta \left(c_k - \sum_{i=1}^m a_{ik} c_i \right). \quad (2.5)$$

Нехай

$$d_j = \sum_{i=1}^m a_{ik} c_i - c_j. \quad (2.6)$$

Значення лінійної форми в новій вершині багатокутника можна знайти з рівняння

$$L_1(\vec{X}) = L(\vec{X}) - \theta d_j. \quad (2.7)$$

Величину d_j називають *оцінкою плану*. В симплексному методі параметри d_j відіграють важливу роль: їх знаки дозволяють визначити чи є опорний план оптимальним. Якщо $d_j > 0$ для всіх j , то даний опорний план є оптимальним, оскільки, на підставі (2.7) і, зважаючи на $\theta \geq 0$, перехід до будь-якої нової вершини веде до спадання лінійної форми. Якщо опорний план неоптимальний, то можливі два випадки:

1. Є хоча б один індекс $j = k$, для якого $d_k < 0$ і всі відповідні компоненти

$$a_{ik} \leq 0 \quad (i = \overline{1, m}).$$

У цьому випадку лінійна форма необмежена зверху і задача немає розв'язку.

2. Для деяких j $d_j < 0$ і для кожного такого j , принаймні, одна з проекцій $a_{ij} > 0$. При переході до наступної вершини лінійна форма зростає і план поліпшується. Для найшвидшого зростання L необхідно в базис включити той вектор \vec{A}_k , для якого оцінка $d_k < 0$ і максимальна по модулю, а вектор \vec{A}_r , для якого значення позитивне і мінімальне, виключити.

Також при лінійному програмуванні використовують методи еліпсоїдів, метод внутрішніх крапок, методи логарифмічних бар'єрних функцій нелінійного програмування. У цих методах вирішення задач лінійного

програмування здійснюється шляхом пошуку в просторі змінних задачі уздовж траєкторій, що не проходять через вершини багатокутника.

Задачам лінійного програмування властиві такі особливості [20]:

1. Цільова функція є зваженою лінійною сумою від невідомих змінних x_i вигляду:

$$J = \sum_{i=1}^n c_i x_i = \max(\min), \quad (2.8)$$

де c_i – коефіцієнти цільової функції. Таку цільову функцію часто називають лінійною формою.

2. Обмеження, що накладаються на область можливих рішень, мають вид лінійної рівності або нерівності:

$$\sum_{j=1}^k a_{ij} x_j \leq b_i, \quad i = 1, 2, \dots, k, \quad (2.9)$$

де a_{ij} , b_i – значення показників цільової функції, причому величини a_{ij} , x_j , b_i позитивні.

В аспекті вирішення задач лінійного програмування розглянемо математичне формулювання і алгоритм розв'язання транспортної задачі.

Змістовна постановка задачі: однорідний продукт, зосереджений у m пунктах відправлення в кількостях a_1, a_2, \dots, a_m одиниць відповідно, необхідно доставити в кожний із n пунктів призначення в кількості b_1, b_2, \dots, b_n одиниць відповідно. Вартість (відстань) перевезення одиниці продукту з i -го пункту відправлення в j -й пункт призначення дорівнює c_{ij} і відома для кожного маршруту.

Нехай x_{ij} – кількість продукту перевезеного з i -го пункту відправлення в j -й пункт призначення. Задача полягає у визначенні таких розмірів x_{ij} для всіх маршрутів, при яких сумарна вартість (відстань) перевезень була б мінімальною.

Математична модель задачі. Позначимо:

- c_{ij} – тарифи (вартість, час, відстань) перевезення одиниці вантажу з i -го пункту відправлення в j -й пункт призначення;

- a_i – запаси вантажу в i -му пункті відправлення;

- b_i – потреба у вантажі в j -му пункті призначення;

- x_{ij} – кількість одиниць вантажу, перевезеного з i -го пункту відправлення в j -й пункт призначення.

Тоді математична модель транспортної задачі про планування перевезень має такий вигляд:

$$y = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min_{x_{ij} \in \Omega}; \quad (2.10)$$

$$\Omega: f_j = \sum_{i=1}^m x_{ij} = b_j, \quad j = \overline{1, n}; \quad (2.11)$$

$$f_{n+i} = \sum_{j=1}^n x_{ij} = a_i, \quad i = \overline{1, m}; \quad (2.12)$$

$$x_{ij} \geq 0; \quad i = \overline{1, m}; \quad j = \overline{1, n}. \quad (2.13)$$

Тут (2.10) – цільова функція, що визначає вартість перевезень усього вантажу. Саме екстремальне (мінімальне) значення цієї функції необхідно знайти в задачі. Причому значення змінних x_{ij} , при яких цільова функція досягає свого мінімуму, повинні належати області припустимих рішень Ω .

Вирази (2.11) – (2.13) визначають область припустимих рішень Ω . При цьому вираз (2.11) відбиває потреби у вантажі в пунктах призначення, вираз (2.12) визначає запаси вантажів у пунктах відправлення, а вираз (2.13) відокремлює від'ємну область значень x_{ij} , в яку дані змінні не можуть потрапляти за своїм фізичним змістом.

Вирази (2.11) – (2.13) називаються обмеженнями задачі лінійного програмування. Вирішення задачі (частковий набір значень змінних x_{ij}) називається припустимим, якщо воно одночасно задовольняє всім обмеженням задачі. Вирішення задачі називається оптимальним, якщо воно забезпечує оптимум (у даному випадку мінімум) функції цілі.

Вважатимемо, що функції y, f_1, f_2, \dots, f_n – неперервні лінійні функції, задані на евклідовому просторі R^n . Дані функції мають місце, коли перевезений вантаж характеризується параметрами, що має вагу, довжину (погонні метри), площу (квадратні метри), об'єм і т.п.

Якщо загальна потреба у вантажі в пунктах призначення дорівнює запасу вантажу в пунктах відправлення, тобто

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j, \quad (2.14)$$

то модель такої транспортної задачі називається *закритою* [21].

У протилежному випадку – *відкритою* [21].

Теорема 2.1 Для розв'язання транспортної задачі необхідно і достатньо, щоб запаси вантажу в пунктах відправлення були рівні потребам у вантажі в пунктах призначення, тобто щоб виконувалася рівність (2.14).

У випадку перевищення запасу над потребою, тобто

$$\sum_{i=1}^m a_i > \sum_{j=1}^n b_j ,$$

вводиться фіктивний $(n+1)$ -й пункт призначення з потребою

$$b_{n+1} = \sum_{i=1}^m a_i - \sum_{j=1}^n b_j .$$

При цьому відповідні тарифи вважаються рівними нулю:

$$c_{i,n+1} = 0 \quad (i = 1, m).$$

Така задача буде вже транспортною задачею, для якої умова (2.14) виконується.

Аналогічно, якщо

$$\sum_{i=1}^m a_i < \sum_{j=1}^n b_j ,$$

вводиться фіктивний $(m+1)$ -й пункт відправлення з запасом вантажу

$$a_{m+1} = \sum_{j=1}^n b_j - \sum_{i=1}^m a_i .$$

При цьому відповідні тарифи вважаються рівними нулю:

$$c_{m+1,j} = 0 \quad (j = 1, n).$$

Така задача буде вже транспортною задачею, для якої умова (2.14) виконується.

Далі будемо розглядати закриту модель транспортної задачі. Якщо ж модель конкретної задачі є відкритою, то, виходячи зі сказаного вище, її необхідно перетворити так, щоб виконувалася рівність (2.14).

У відкритій моделі область припустимих значень (за інших рівних умов) значно ширше, тому цільова функція досягає кращих значень або, принаймні, не гірших.

Особливості вирішення закритої транспортної задачі.

Визначення 2.1 Усяке додатне рішення систем лінійних рівнянь (2.11) і (2.12), що обумовлене матрицею $X = [x_{ij}]$, $i = 1, m$, $j = 1, n$, називається планом транспортної задачі.

Визначення 2.2 План $X^* = [x^*_{ij}]$, $i = 1, m$, $j = 1, n$, при якому функція (2.10) приймає своє мінімальне значення, називається оптимальним планом транспортної задачі.

Число змінних x_{ij} транспортної задачі з m пунктами відправлення і n пунктами призначення дорівнює mn , а число рівнянь у системах (2.11) і (2.12) дорівнює $m+n$. Оскільки передбачається, що виконується умова (2.14), то число лінійно незалежних рівнянь дорівнює $m+n-1$. Отже, опорний план транспортної задачі може мати не більше $m+n-1$ відмінних від нуля невідомих.

Визначення 2.3 План $X^* = [x^*_{ij}]$, $i = 1, m$, $j = 1, n$ є опорним не виродженим, якщо в ньому кількість відмінних від нуля компонентів дорівнює $m+n-1$, а якщо менше – то виродженим.

Для визначення опорного плану існує декілька методів. Один з них – метод північно-західного кута (буде розглянутий нижче).

Як і для всякої задачі лінійного програмування, оптимальний план транспортної задачі є також опорним планом. Для визначення оптимального плану транспортної задачі можна використати диференціальний алгоритм, симплекс-метод та інші універсальні методи. Однак, через виняткову практичну важливість цієї задачі і специфіку її обмежень (кожна невідома входить лише в два рівняння систем (2.11) і (2.12), а коефіцієнти при невідомих дорівнюють одиниці) для визначення оптимального плану транспортної задачі розроблені спеціальні методи. Один з них – метод потенціалів.

Вихідні дані транспортної задачі записують у вигляді таблиці 2.1.

Таблиця 2.1 – Вихідні дані транспортної задачі

Пункти відправлення	Запаси	Пункти призначення					
		1	2	...	j	...	n
		Потреби					
		b_1	b_2	...	b_j	...	b_n
1	a_1	$c_{11} \quad x_{11}$	$c_{12} \quad x_{12}$...	$c_{1j} \quad x_{1j}$...	$c_{1n} \quad x_{1n}$
2	a_2	$c_{21} \quad x_{21}$	$c_{22} \quad x_{22}$...	$c_{2j} \quad x_{2j}$...	$c_{2n} \quad x_{2n}$
...
i	a_i	$c_{i1} \quad x_{i1}$	$c_{i2} \quad x_{i2}$...	$c_{ij} \quad x_{ij}$...	$c_{in} \quad x_{in}$
...
m	a_m	$c_{m1} \quad x_{m1}$	$c_{m2} \quad x_{m2}$...	$c_{mj} \quad x_{mj}$...	$c_{mn} \quad x_{mn}$

Визначення початкового опорного плану транспортної задачі.

Вирішення транспортної задачі починають із знаходження будь-якого опорного плану. Для цього розроблені специфічні методи. Один з них одержав у назву «метод північно-західного кута», іноді його називають «діагональним методом» або «методом перехідних приступів» [22].

Сутність методу полягає в тому, що опорний план знаходять за $m+n-1$ кроками, на кожному з яких у таблиці транспортної задачі заповнюють одну клітинку. Заповнення однієї клітинки забезпечує повне або часткове задоволення потреби у вантажі одного з пунктів призначення (відповідно до того, в стовпці якого знаходиться клітинка), або вивіз вантажу з одного з пунктів відправлення (відповідно з того, в рядку якого знаходиться клітинка).

Заповнення таблиці слід починати з лівого верхнього квадрату (північно-західного кута). З позиції цього квадрату порівнюють запас вантажу в першому пункті відправлення з потребою першого пункту призначення. Вибирають менший розмір і записують у даний квадрат, який з цього моменту стає «зайнятим». Якщо в клітинку записується потреба пункту призначення, то з подальшого розгляду виключають відповідний стовпець таблиці і переходять у ліву сусідню клітинку. Якщо в клітинку записується запас пункту відправлення, то з подальшого розгляду виключають відповідний рядок таблиці і переходять у сусідню клітинку, що знаходиться нижче заповненої.

У новій клітинці для частини таблиці, що залишилася, повторюють процедуру першого кроку з урахуванням зміни запасу вантажу одного з відправників або потреби у вантажі одного з одержувачів у результаті попереднього кроку.

Після $m+n-2$ описаних вище кроків одержують задачу з одним пунктом відправлення і одним пунктом призначення. При цьому залишається вільною тільки одна клітинка, а запаси пункту відправлення дорівнюватимуть потребам пункту призначення. Заповнення цієї клітинки, тобто здійснення $(m+n-1)$ -го кроку приводить до шуканого опорного плану.

Слід зауважити, що в процесі використання методу північно-західного кута може трапитися, що потреби у вантажі чергового пункту призначення рівні запасам чергового пункту відправлення. У цьому випадку з подальшого розгляду виключають або стовпець, або рядок, тобто тільки що-небудь одне. Таким чином, запаси відповідного пункту відправлення або потребу відповідного пункту призначення вважають рівними нулю. Цей нуль записують у чергову клітинку, яка заповнюється [4].

Зазначені вище умови гарантують одержання $m+n-1$ зайнятих клітинок, у яких знаходяться компоненти опорного плану.

Опорний план перевезень повинен відповідати таким вимогам:

- по-перше, кількість зайнятих маршрутів (клітинок) повинно бути на одиницю менше суми числа постачальників і числа споживачів, тобто $m+n-1$;
- по-друге, не повинно бути жодного зайнятого маршруту (клітинки), що опинився б єдиним і в рядку, і в стовпці таблиці.

Визначення оптимального опорного плану транспортної задачі [23].

Для визначення оптимального плану транспортної задачі розроблено декілька методів. Найбільш часто використовується *метод потенціалів*. Метод припускає, що вже відомий якийсь опорний план. Його можна одержати, наприклад, розглянутим методом північно-західного кута. Вихідний опорний план необхідно перевірити на оптимальність.

Теорема 2.2 Якщо для деякого опорного плану

$$X^* = [x_{ij}^*], i = 1, m, j = 1, n$$

транспортної задачі з заданими тарифами перевезень c_{ij} існують такі числа

$$\alpha_i (i = 1, m) \text{ і } \beta_j (j = 1, n),$$

$$\text{що} \quad \beta_i - \alpha_j = c_{ij} \text{ при } x_{ij} > 0 \quad (2.15)$$

$$\text{і} \quad \beta_i - \alpha_j \leq c_{ij} \text{ при } x_{ij} = 0 \quad (2.16)$$

для всіх $i=1, m$ і $j=1, n$, то $X^* = [x_{ij}^*]$ – оптимальний план.

Визначення 2.4 Числа α_i ($i = 1, m$) і β_j ($j = 1, n$) називаються потенціалами відповідно пунктів відправлення і пунктів призначення.

Теорема 2.2 дозволяє побудувати алгоритм знаходження рішення транспортної задачі.

Нехай знайдений опорний план транспортної задачі. Для кожного з пунктів відправлення і призначення визначають потенціали α_i ($i = 1, m$) і β_j ($j = 1, n$) із системи рівнянь:

$$\beta_i - \alpha_j = c_{ij}. \quad (2.17)$$

Так як число заповнених клітинок дорівнює $n+m-1$, то система (2.17) із $n+m$ невідомими містить $n+m-1$ рівнянь. Оскільки число невідомих перевищує на одиницю число рівнянь, одне з невідомих потрібно взяти рівним

довільному числу, наприклад, $\alpha_1 = 0$, і знайти послідовно із системи (2.17) значення інших невідомих.

Після того, як усі потенціали знайдені, дія кожної з вільних клітинок визначають числа

$$\alpha_{ij} = \beta_i - \alpha_j - c_{ij}.$$

Якщо серед чисел α_{ij} немає додатних, то знайдений опорний план є оптимальним. Якщо ж для деякої вільної клітинки $\alpha_{ij} > 0$, то опорний план, що перевіряється, не є оптимальним, і треба перейти до нового опорного плану. Для цього розглядають усі вільні клітинки, для яких $\alpha_{ij} > 0$, і вибирають ту, для якої число α_{ij} максимальне. Обрану клітку необхідно заповнити.

Заповнюючи обрану клітинку, треба змінити обсяги перевезень, записаних у ряді інших зайнятих клітинках і пов'язаних з обраним циклом.

Визначення 2.5 Циклом у таблиці транспортної задачі *називається замкнута ломана лінія*, вершини якої розташовані в зайнятих клітинках таблиці, а ланки – уздовж рядків і стовпців, причому в кожній вершині циклу зустрічаються рівно дві ланки, одна з яких знаходиться в рядку, а інша – у стовпці.

Якщо ломана лінія, що складає цикл, перетинається сама із собою, то точка самоперетину не є вершиною.

При правильній побудові опорного плану для будь-якої вільної клітинки можна побудувати тільки один цикл. Після того як для обраної вільної клітинки він побудований, необхідно перейти до нового опорного плану. Для цього треба перемістити вантажі в межах клітинок, що утворюють цикл. Переміщення роблять за такими правилами:

- кожній з кліток, пов'язаних циклом з обраною вільною клітинкою, приписують знак «+» або «-», причому вільній клітинці – знак плюс, а всім іншим клітинкам – по черзі знаки мінус і плюс;

- у вільну клітинку переносять менше з чисел x_{ij} , що знаходиться в «мінусових» клітинках, і одночасно це число додають до відповідних чисел, що знаходяться в «плюсових» клітинках, і віднімають із чисел, що знаходяться в «мінусових» клітинках. Клітинка, що раніше була вільною стає зайнятою, а «мінусова» клітинка, в якій стояло мінімальне число x_{ij} стає вільною.

У результаті зазначених вище переміщень вантажів у межах клітинок, пов'язаних циклом з обраною вільною клітинкою, визначають новий опорний план транспортної задачі. Число зайнятих клітинок залишається рівним $n+m-1$. Якщо в зайнятих «мінусових» клітинках циклу є два і більше однакових мінімальних чисел x_{ij} , то звільняють тільки одну з таких клітинок, а інші залишають зайнятими з нульовими поставаннями.

Отриманий новий опорний план транспортної задачі перевіряють на оптимальність. Для цього визначають потенціали пунктів відправлення і призначення і знаходять числа

$$\alpha_{ij} = \beta_i - \alpha_j - c_{ij}$$

для всіх вільних клітинок. Якщо серед цих чисел не буде додатних, то це означає, що новий опорний план є оптимальним. Якщо ж є додатні числа, то необхідно перейти до нового опорного плану. У результаті ітераційного плану кінцевого числа переходів одержують оптимальний план процесу задачі.

Таким чином, процес знаходження вирішення транспортної задачі методом потенціалів включає наступні етапи [24]:

1-й етап. Знаходять опорний план.

2 й етап. Знаходять потенціали пунктів відправлення і призначення.

3-й етап. Визначають числа α_{ij} для кожної вільної клітинки. Якщо серед них немає додатних, то отримано оптимальний план транспортної задачі, у противному разі переходять до нового опорного плану.

4-й етап. Вибирають серед додатних чисел α_{ij} максимальне, будують для відповідної вільної клітинки цикл перерахування і роблять зсув за циклом, одержуючи при цьому новий опорний план. Далі переходять до 2-го етапу.

Розглянемо приклад вирішення транспортної задачі методом потенціалів.

Розглянемо приклад. Три заводи, що виготовляють бетонні конструкції, постачаються цементом з чотирьох складів. Попит заводів b_j відповідно дорівнює 280, 90 і 180 тис. т/міс. Пропускна здатність складів a_i відповідно становить 200, 150, 80 і 120 тис. т/міс. Відстані перевезень із i -го складу на j -й завод подані в матриці (км)

$$C = [c_{ij}] = \begin{pmatrix} 1 & 5 & 3 \\ 6 & 8 & 9 \\ 2 & 7 & 4 \end{pmatrix}.$$

Потрібно скласти план перевезень цементу зі складів на заводи, що задовольняв би пропускним спроможностям складів і потребам заводу, а сумарний пробіг вантажного транспорту був би мінімальним.

Розв'язання

Позначимо через x_{ij} – кількість цементу, який щомісяця потрібно доставляти на j -й завод із i -го складу.

Тоді математична модель задачі має вигляд:

$$y = x_{11} + 5x_{12} + 3x_{13} + 6x_{21} + 8x_{22} + 9x_{23} + 2x_{31} + 7x_{32} + 4x_{33} + 4x_{41} + x_{42} + 11x_{43} \rightarrow \min; \quad (2.18)$$

$$x_{ij} \in \Omega;$$

$$\Omega: \quad x_{11} + x_{21} + x_{31} + x_{41} = 280; \quad (2.19)$$

$$x_{12} + x_{22} + x_{32} + x_{42} = 90; \quad (2.20)$$

$$x_{13} + x_{23} + x_{33} + x_{43} = 180; \quad (2.21)$$

$$x_{11} + x_{12} + x_{13} = 200; \quad (2.22)$$

$$x_{21} + x_{22} + x_{23} = 150; \quad (2.23)$$

$$x_{31} + x_{32} + x_{33} = 80; \quad (2.24)$$

$$x_{41} + x_{42} + x_{43} = 120; \quad (2.25)$$

$$x_{ij} \geq 0, \quad (2.26)$$

де (2.18) – цільова функція;

(2.19) – (2.21) – обмеження задачі, що визначають місячні запаси цементу на складах;

(2.22) – (2.25) – обмеження задачі, що визначають місячну потребу в цементі на заводах;

(2.26) – обмеження, що визначає неможливість від'ємних значень для постачань цементу на заводи.

1 крок.

1-й етап. Використовуючи метод північно-західного кута, знайдемо опорне рішення транспортної задачі (2.18) – (2.26).

Відповідно до цього методу заповнюємо таблицю, починаючи з лівого верхнього квадрату. Порівнюємо запас вантажу в першому пункті відправлення (200 тис. т/міс.) із потребою першого пункту призначення (280 тис. т/міс.).

Вибираємо меншу величину (200) і записуємо її в даний квадрат. Оскільки весь запас у першому пункті відправлення вичерпаний, то з

подальшого розгляду виключаємо перший рядок і переходимо в сусідню клітинку, що знаходиться нижче заповненої. У новій клітинці для частини таблиці, що залишилася, повторюємо процедуру заповнення верхньої лівої клітинки, але з урахуванням того, що потреба першого пункту призначення зменшилася на 200 тис. т/міс. і стала рівною 80 тис. т/міс. Тобто порівнюємо запас другого пункту відправлення (150 тис. т/міс.) із новою потребою першого пункту призначення (80 тис. т/міс). Вибираємо меншу величину (80) і записуємо її в нову клітинку.

Оскільки потреба у вантажі в першому пункті призначення повністю задоволена, то з подальшого розгляду виключаємо перший стовпець і переходимо в сусідню клітинку, що знаходиться справа від тільки що заповненої. Для нової верхньої лівої клітинки частини таблиці, що залишилася, повторюємо процедуру заповнення з урахуванням зміни запасу в другому пункті відправлення на 50 тис. т/міс.

І так доти, поки не буде заповнено $m+n-2$ клітинок.

Остання $(m+n-2)$ -я клітинка заповнюється механічно – у неї записується залишкова потреба останнього пункту призначення або залишковий запас останнього пункту відправлення. В умовах задачі це величина 120.

Усі проміжні результати по знаходженню початкового опорного плану:

$$X_0 = \begin{pmatrix} 200 & 0 & 0 \\ 80 & 70 & 0 \\ 0 & 20 & 60 \\ 0 & 0 & 120 \end{pmatrix}$$

відображені у таблиці 2.2.

Для початкового опорного плану обчислюємо значення (2.18) цільової функції:

$$y_0 = 1 \cdot 200 + 6 \cdot 80 + 8 \cdot 70 + 7 \cdot 20 + 4 \cdot 60 + 11 \cdot 120 = 2940 \text{ тис. т/міс.}$$

Це значення буде використано на наступних кроках для контролю просування до оптимуму. Значення цільової функції повинно послідовно зменшуватися з кожним кроком.

Таблиця 2.2 – Проміжні результати по знаходженню початкового опорного плану

Пункт відправлення	Запас вантажу	Пункт призначення			Потенціал пункту відправлення α_i
		1	2	3	
		Потреба			
		280	90	180	
1	200	1 200	5	3	0
2	150	6 80	8 70	9	-5
3	80	2	7 - 20	4 + 60	-4
4	120	4	1 +	11 - 120	-11
Потенціал пункту призначення β_i		1	3	0	

2-й етап. Знайдений опорний план перевіряємо на оптимальність. Обчислюємо потенціали пунктів відправлення і призначення із системи:

$$\begin{aligned} \beta_1 - \alpha_1 &= 1, & \beta_2 - \alpha_2 &= 8, & \beta_3 - \alpha_3 &= 4, \\ \beta_1 - \alpha_2 &= 6, & \beta_2 - \alpha_3 &= 7, & \beta_3 - \alpha_4 &= 11, \end{aligned}$$

що містить шість рівнянь із сімома невідомими. Вважаючи $\alpha_1=0$, знаходимо:

$$\beta_1=1, \alpha_2=-5, \beta_2=3, \alpha_3=-4, \beta_3=0, \alpha_4=-11.$$

Записуємо знайдені потенціали в таблицю 2.2.

3-й етап. Для кожної вільної клітинки обчислюємо числа $\alpha_{ij}=\beta_i-\alpha_j-c_{ij}$:

$$\alpha_{12}=-2, \alpha_{13}=-3, \alpha_{23}=-4, \alpha_{31}=3, \alpha_{41}=8, \alpha_{42}=13.$$

Записуємо знайдені числа у відповідні вільні клітинки таблиці 2.2 і розміщуємо їх у рамочки, щоб відрізнити їх від іншої інформації в таблиці. Так як серед чисел α_{ij} є додатні, то опорний план X_0 не є оптимальним.

4-й етап. Серед додатних чисел α_{ij} вибираємо максимальне: $\alpha_{42}=13$. Для відповідної вільної клітинки будуємо цикл, а саму клітинку позначаємо знаком «+». У таблиці 2.2 зайняті клітинки, що складають цикл, виділені сірим фоном. Потім позначаємо знаками «-» і «+» по черзі інші клітинки циклу, слідуючи уздовж ломаної лінії циклу.

Найменшим із чисел x_{ij} у «мінусових» клітинках є x_{32} (20). Дана клітинка стає вільною, а інші клітинки циклу змінюють свої значення в такий спосіб:

$$x_{42}=20, x_{43}=120-20=100, x_{33}=60+20=80.$$

У результаті зроблених перетворень одержуємо новий опорний план:

$$X_1 = \begin{pmatrix} 200 & 0 & 0 \\ 80 & 70 & 0 \\ 0 & 0 & 80 \\ 0 & 20 & 100 \end{pmatrix}.$$

При такому опорному плані функція цілі (2.18) стає рівною 2680 тис. т/міс, що менше вихідного значення 2940 тис. т/міс.

На цьому закінчується 1-й крок оптимізації. На наступному кроці процедура 1-го кроку повторюється, але без 1-го етапу.

2-й крок.

Аналізуємо новий опорний план (табл. 2.3) на оптимальність. Знову знаходимо потенціали пунктів відправлення і пунктів призначення, для чого складаємо таку систему рівнянь:

$$\begin{aligned} \beta_1 - \alpha_1 &= 1, & \beta_2 - \alpha_2 &= 8, & \beta_3 - \alpha_4 &= 1, \\ \beta_1 - \alpha_2 &= 6, & \beta_3 - \alpha_3 &= 4, & \beta_3 - \alpha_4 &= 11. \end{aligned}$$

Вважаючи $\alpha_1 = 0$, знаходимо

$$\beta_1 = 1, \alpha_2 = -5, \beta_2 = 3, \alpha_4 = 2, \beta_3 = 13, \alpha_3 = 9.$$

Для кожної вільної клітинки обчислюємо числа α_{ij} :

$$\alpha_{12} = -2, \alpha_{13} = 10, \alpha_{23} = 9, \alpha_{31} = -10, \alpha_{32} = -13, \alpha_{41} = -5.$$

Якщо серед чисел α_{ij} є додатні ($\alpha_{13}=10$, $\alpha_{23}=9$), то опорний план X_1 не є оптимальним.

Таблиця 2.3 – Новий опорний план

Пункт відправлення	Запас вантажу	Пункт призначення			Потенціал пункту відправлення α_i
		1	2	3	
		Потреба			
		280	90	180	
1	200	1 - 200	5	3 +	0
2	150	6 + 80	8 - 70	9	-5
3	80	2	7	4 80	9
4	120	4	1 + 20	11 - 100	2
Потенціал пункту призначення β_i		1	3	13	

Серед додатних чисел α_{ij} вибираємо максимальне: $\alpha_{13}=10$. Для відповідної вільної клітинки будуємо цикл, а саму клітинку позначає знаком «+».

Позначаємо вузлові клітинки циклу по черзі знаками «-» і «+».

Найменшим із чисел x_{ij} у «мінусових» клітинках є x_{23} (70). Дана клітинка стає вільною, а інші клітинки циклу змінюють свої значення в так спосіб:

$$x_{11}=200-70=130, x_{13}=70, x_{21}=80+70=150, x_{42}=20+70=90, x_{43}=100-70=30.$$

У результаті виконаних перетворень одержуємо новий опорний план:

$$X_2 = \begin{pmatrix} 130 & 0 & 70 \\ 150 & 0 & 0 \\ 0 & 0 & 80 \\ 0 & 90 & 30 \end{pmatrix}.$$

При такому опорному плані функція (2.18) стає рівною 1980 тис. т/міс, що значно менше попереднього значення 2680 тис. т/міс.

3-й крок.

Аналізуємо новий опорний план (табл. 2.4) на оптимальність. Знову знаходимо потенціали пунктів відправлення і пунктів призначення, для чого складаємо таку систему рівнянь:

$$\begin{array}{lll} \beta_1 - \alpha_1 = 1, & \beta_1 - \alpha_2 = 6, & \beta_2 - \alpha_4 = 1, \\ \beta_3 - \alpha_1 = 3, & \beta_3 - \alpha_3 = 4, & \beta_3 - \alpha_4 = 11. \end{array}$$

Вважаючи $\alpha_1 = 0$, знаходимо

$$\beta_1 = 1, \beta_3 = 3, \alpha_2 = -5, \beta_3 = 4, \alpha_3 = 0, \alpha_4 = -8, \beta_2 = -7.$$

Для кожної вільної клітинки обчислюємо числа α_{ij} :

$$\alpha_{12} = -12, \alpha_{22} = -10, \alpha_{23} = -1, \alpha_{31} = -1, \alpha_{32} = -14, \alpha_{41} = 5.$$

Оскільки серед чисел α_{ij} одне додатне ($\alpha_{41}=5$), то опорний план X_2 не є оптимальним.

Таблиця 2.4 – Новий опорний план

Пункт Відправлення	Запас вантажу	Пункт призначення			Потенціал пункту відправлення α_i
		1	2	3	
		Потреба			
		280	90	180	
1	200	1 - 130	5	3 + 70	0
2	150	6 150	8	9	-5
3	80	2	7	4 80	0
4	120	4 +	1 90	11 - 30	-8
Потенціал пункту призначення β_i		1	-7	3	

Для відповідної вільної клітинки (нижньої лівої) будуємо цикл, а саму клітинку позначаємо знаком «+». У таблиці 2.4 позначаємо вузлові клітинки циклу по черзі знаками «-» і «+». Найменшим із чисел x_{ij} у «мінусових» клітинках є x_{43} (30). Дана клітинка стає вільною, а інші клітинки циклу змінюють свої значення в такий спосіб:

$$x_{11}=130-30=100, x_{13}=70+30=100, x_{41}=30.$$

У результаті зроблених перетворень одержуємо новий опорний план:

$$X_3 = \begin{pmatrix} 100 & 0 & 100 \\ 150 & 0 & 0 \\ 0 & 0 & 80 \\ 30 & 90 & 0 \end{pmatrix}.$$

При такому опорному плані функція (2.18) стає рівною 1830 тис. т/міс, що менше попереднього значення 1980 тис. т/міс.

4-й крок.

Аналізуємо новий опорний план (табл. 2.5) на оптимальність. Знову знаходимо потенціали пунктів відправлення і пунктів призначення, для чого складаємо таку систему рівнянь:

$$\begin{array}{lll} \beta_1 - \alpha_1 = 1, & \beta_1 - \alpha_2 = 6, & \beta_1 - \alpha_4 = 4, \\ \beta_3 - \alpha_1 = 3, & \beta_3 - \alpha_3 = 4, & \beta_2 - \alpha_4 = 1. \end{array}$$

Вважаючи $\alpha_1 = 0$, знаходимо

$$\beta_1=1, \beta_3=3, \alpha_2=-5, \alpha_3=-1, \alpha_4=-3, \beta_2=-2.$$

Для кожної вільної клітки обчислюємо числа α_{ij} :

$$\alpha_{12}=-7, \alpha_{22}=-4, \alpha_{23}=-1, \alpha_{31}=0, \alpha_{32}=-8, \alpha_{41}=-5.$$

Так як серед чисел α_{ij} немає строго додатних, то опорний план X_3 є оптимальним.

Таблиця 2.5 – Новий опорний план

Пункт відправлення	Запас вантажу	Пункт призначення			Потенціал пункту відправлення α_i
		1	2	3	
		Потреба			
		280	90	180	
1	200	1100	5	3100	0
2	150	6150	8	9	-5
3	80	2	7	480	-1
4	120	430	190	11	-3
Потенціал пункту призначення β_i		1	-2	3	

Різновиди транспортних задач.

Розглянута математична модель (2.10) – (2.13) є класичною моделлю транспортної задачі [23]. У реальній практиці, як правило, транспортна задача зустрічається в дещо іншій постановці. Математична модель реальної транспортної задачі може відрізнятися від класичної або видом цільової функції, або видом обмежень, або характером змінних, або будь-яким сполученням перерахованих відмінностей одночасно.

Розглянемо декілька модифікацій транспортної задачі.

Транспортна задача про розподіл випуску продукції.

При комплексному вирішенні проблеми виробництва і реалізації продукції виникає задача, що полягає у визначенні такого плану випуск і перевезень готової продукції, при якому досягаються мінімальні витрати на її виготовлення і доставку споживачам.

Для вирішення даної задачі розглядається повна собівартість виробництва одиниці продукції на кожному підприємстві (s_i) і транспортні витрати (s_{ij}), що залежать від типу застосованих транспортних засобів і районів розташування заводів-виробників і споживачів.

Математична модель такої задачі має вигляд:

$$y = \sum_{i=1}^m \sum_{j=1}^n (s_i + s_{ij}) x_{ij} \rightarrow \min; \quad x_{ij} \in \Omega \quad (2.27)$$

$$\Omega: \quad f_j = \sum_{i=1}^m x_{ij} = b_j, \quad j = \overline{1, n}; \quad (2.28)$$

$$f_{n+i} = \sum_{j=1}^n x_{ij} = a_i, \quad i = \overline{1, m}; \quad j = \overline{1, n}. \quad (2.29)$$

$$x_{ij} \geq 0; \quad i = \overline{1, m}; \quad j = \overline{1, n}. \quad (2.30)$$

$$x_{ij} = \text{int}; \quad i = \overline{1, m}; \quad j = \overline{1, n}. \quad (2.31)$$

Якщо за умовою задачі потрібні ще капітальні вкладення в засоби транспорту, то показником ефективності служать приведені витрати, (2.27) матиме вигляд:

$$y = \sum_{i=1}^m \sum_{j=1}^n (s_i + s_y + E_n k_i) \cdot x_{ij} \rightarrow \min, \quad (2.32)$$

де E_n – нормативний коефіцієнт ефективності капітальних вкладень;

k – питомі капітальні вкладення, що приводяться на одиницю перевезень.

Виконуючи підстановку

$$C_{ij} = s_i + s_{ij}$$

в цільову функцію (2.27), і підстановку

$$C_{ij} = s_i + s_{ij} + E_n k_i$$

в цільову функцію (2.32), задача (2.27) – (2.31) і задача (2.32), (2.28) – (2.31) приводяться до класичної транспортної задачі, що може бути вирішена методом потенціалів.

Розподільна транспортна задача про вибір засобів доставки вантажу.

Змістовна постановка задачі. Нехай через $j = 1, 2, \dots, n$ позначено пункти, що мають вантажі для перевезень об'ємами a_j відповідно. Є m засобів доставки вантажу (видів транспорту). Вантажопідйомність i -го засобу доставки складає p_i , а наявний його парк дорівнює b_i , $i = 1, 2, \dots, m$. Вантажі підлягають доставці в один центральний пункт (склад). Витрати при здійсненні однією одиницею i -го засобу доставки від j -го пункту до складу дорівнюють c_{ij} . Потрібно скласти найбільш оптимальний план доставки.

Математична модель задачі. Позначимо через x_{ij} кількість засобів доставки i -го типу, що відправляється з j -го пункту. Тоді математична модель розподільної транспортної задачі про вибір транспортних засобів має вигляд:

$$y = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min ; \quad (2.33)$$

$$\Omega : \sum_{i=1}^m p_i x_{ij} \geq a_j, j = \overline{1, n} ; \quad (2.34)$$

$$\sum_{j=1}^n x_{ij} \leq b_i, i = \overline{1, m} ; \quad (2.35)$$

$$x_{ij} \geq 0; i = \overline{1, m}; j = \overline{1, n} ; \quad (2.36)$$

$$x_{ij} = \text{int}; i = \overline{1, m}; j = \overline{1, n} . \quad (2.37)$$

Цільова функція (2.33) визначає сумарні витрати доставки вантажу на центральний склад. Вираз (2.34) вказує на необхідність вивезення всього вантажу з пунктів відправлення. Обмеження (2.35) вказує на те, що кількість використаних засобів доставки не повинна перевищувати їхній наявний парк.

Поява параметра p_i у системі обмежень (2.34) перешкоджає зведенню математичної моделі до моделі класичної, тому вирішувати її методом потенціалів неможливо. Вирішення даної задачі класичними методами лінійного програмування також неможливе через цілочисельність змінних x_{ij} . Розв'язання задачі можна одержати методом відсікання (шляхом введення в задачу додаткових обмежень у вигляді нерівностей Гоморі, однак, процедура вирішення різко ускладнюється) [24 – 26]. Вирішення задачі найбільш доцільно покласти на програму Solver (пошук рішення) інформаційної системи Microsoft Excel.

Транспортна задача про двоетапне перевезення вантажу.

Змістовна постановка задачі. Однорідний вантаж потрібно доставити з пунктів відправлення в пункти призначення. При доставці в пункти призначення вантажі можуть бути спочатку доставлені на p перевалочних пунктів. Задано вартості перевезень c_y з кожного пункту відправлення в кожний пункт призначення і перевалочний пункт, а також вартості перевезення з кожного перевалочного пункту в пункт призначення.

Математична модель завдання. Позначимо:

- c_{ij} – вартість перевезення одиниці вантажу з i -го пункту відправлення в j -й пункт призначення, $i = \overline{1, v}$, $j = \overline{1, n}$;
- c_{ik} – вартість перевезення одиниці вантажу з i -го пункту відправлення в k -й перевалочний пункт, $i = \overline{1, m}$, $k = \overline{1, p}$;
- c_{kj} – вартість перевезення одиниці вантажу з k -го перевалочного пункту в j -й пункт призначення, $k = \overline{1, p}$, $j = \overline{1, n}$;
- a_i – запаси вантажу в i -му пункті відправлення;
- b_j – потреба у вантажі в j -му пункті призначення;
- c_k – місткість k -го перевалочного пункту;
- x_{ij} – кількість вантажу перевезеного з i -го пункту відправлення в j -й пункт призначення;
- y_{ik} – кількість вантажу перевезеного з i -го пункту відправлення в k -й перевалочний пункт;
- z_{kj} – кількість вантажу перевезеного з k -го перевалочного пункту в j -й пункт призначення.

Математична модель задачі, з урахуванням вище наведених позначень, може бути подана у вигляді задачі лінійного програмування:

$$y = \sum_{i=1}^m \sum_{j=1}^n c_{ij} * x_{ij} + \sum_{i=1}^m \sum_{k=1}^p c_{ik} * y_{ik} + \sum_{k=1}^p \sum_{j=1}^n c_{kj} * z_{kj} \rightarrow \min_{x_{ij}, y_{ik}, z_{kj} \in \Omega}; \quad (2.38)$$

$$\Omega: \sum_{j=1}^n x_{ij} + \sum_{k=1}^p y_{ik} \leq a_i, \quad i = \overline{1, m}; \quad (2.39)$$

$$\sum_{i=1}^m x_{ij} + \sum_{k=1}^p z_{kj} \geq p_j, \quad j = \overline{1, n}; \quad (2.40)$$

$$\sum_{i=1}^m y_{ik} \leq c_k, \quad k = \overline{1, p}; \quad (2.41)$$

$$\sum_{i=1}^m y_{ik} = \sum_{j=1}^n z_{kj}, \quad k = \overline{1, p}; \quad (2.42)$$

$$x_{ij} \geq 0; \quad y_{ik} \geq 0; \quad z_{kj} \geq 0, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad k = \overline{1, p}. \quad (2.43)$$

Тут цільова функція (2.38) складається з витрат трьох видів [27]:

- на доставку частини вантажу з пунктів відправлення в пункти призначення, маючи перевалочні пункти;
- на перевезення частини вантажу з пункт призначення в перевалочні пункти;
- на доставку вантажу з перевалочних пунктів у пункти призначення.

Система обмежень (2.39) говорить про те, що сумарні об'єми вантажів, що вивозяться з пунктів відправлення, не можуть перевищувати запаси вантажів у цих пунктах. Система обмежень (2.40) свідчить про те, що сумарні об'єми вантажів, що надходять у пункти призначення, не можуть бути менше відповідних потреб пунктів призначення. Система обмежень (2.41) означає, що сумарне завезення вантажів на кожний перевалочний пункт не може перевищувати його місткості. Система обмежень (2.42) вказує на те, що весь вантаж із перевалочних пунктів повинен бути вивезений повністю.

Як і в попередній задачі, математична модель (2.38) – (2.43) не може бути приведена до класичної. Тому вирішення задачі найбільш доцільно покласти на програму Solver (пошук рішення) інформаційної системи Microsoft Excel.

Транспортна задача про двоетапне перевезення вантажу декількох видів.

Змістовна постановка завдання. Вантаж, що включає q видів продукції, потрібно доставити з пунктів відправлення в пункти призначення. При доставці

в пункти призначення вантажі можуть бути спочатку доставлені на q перевалочних пунктів. Задано вартості перевезень для кожного виду вантажу з кожного пункту відправлення в кожний пункт призначення і перевалочний пункт, а також вартості перевезення з кожного перевалочного пункту в кожний пункт призначення.

Математична модель завдання. Позначимо:

- c_{ijl} – вартість перевезення одиниці l -го виду вантажу з i -го пункту відправлення в j -й пункт призначення, $i = \overline{1, m}$, $j = \overline{1, n}$, $l = \overline{1, q}$;
- c_{ikl} – вартість перевезення одиниці l -го виду вантажу з i -го пункту відправлення в k -й перевалочний пункт, $i = \overline{1, m}$, $k = \overline{1, p}$, $l = \overline{1, q}$;
- c_{kjl} – вартість перевезення одиниці l -го виду вантажу з k -го перевалочного пункту в j -й пункт призначення, $k = \overline{1, p}$, $j = \overline{1, n}$, $l = \overline{1, q}$;
- a_{il} – запаси l -го виду вантажу в i -му пункті відправлення;
- b_{jl} – потреба в l -м виді вантажу в j -му пункті призначення;
- c_{kl} – місткість k -го перевалочного пункту стосовно l -го виду вантажу;
- x_{ijl} – кількість l -го виду вантажу перевезеного з i -го пункту відправлення в j -й пункт призначення;
- y_{ikl} – кількість l -го виду вантажу перевезеного з i -го пункту відправлення в k -й перевалочний пункт;
- z_{kjl} – кількість l -го виду вантажу перевезеного з k -го перевалочного пункту в j -й пункт призначення.

Математична модель задачі, з урахуванням вище приведених позначень, може бути подана у вигляді задачі лінійного програмування:

$$y = \sum_{i=1}^m \sum_{j=1}^n \sum_{l=1}^q c_{ijl} * x_{ijl} + \sum_{i=1}^m \sum_{k=1}^p \sum_{l=1}^q c_{ikl} * y_{ikl} + \sum_{k=1}^p \sum_{j=1}^n \sum_{l=1}^q c_{kjl} * z_{kjl} \rightarrow \min_{x_{ijl}, y_{ikl}, z_{kjl} \in \Omega}; \quad (2.44)$$

$$\Omega: \sum_{j=1}^n x_{ijl} + \sum_{k=1}^p y_{ikl} \leq a_{il}, \quad i = \overline{1, m}; \quad (2.45)$$

$$\sum_{i=1}^m x_{ijl} + \sum_{k=1}^p z_{kjl} \geq b_{jl}, \quad j = \overline{1, n}, \quad j = \overline{1, q}; \quad (2.46)$$

$$\sum_{i=1}^m y_{ikl} \leq c_{kl}, \quad k = \overline{1, p}, \quad l = \overline{1, q}; \quad (2.47)$$

$$\sum_{i=1}^m y_{ik} = \sum_{j=1}^n z_{kj}, \quad k = \overline{1, p}, \quad l = \overline{1, q}; \quad (2.48)$$

$$x_{ijl} \geq 0; \quad y_{kjl} \geq 0; \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad k = \overline{1, p}, \quad l = \overline{1, q}. \quad (2.49)$$

Математична модель задачі відрізняється від попередньої тільки там, що вона враховує різновид вантажів.

Транспортна задача про двоетапне перевезення вантажів декількох видів за запитами споживачів.

Існує модифікація транспортної задачі двоетапного перевезення вантажів декількох видів, у якій кількість вантажу в пунктах відправлення не фіксована. Вона залежить від запитів споживачів.

Математична модель задачі. Позначимо:

- c_{ijl} – вартість перевезення одиниці l -го виду вантажу з i -го пункту відправлення в j -й пункт призначення, $i = \overline{1, m}, \quad j = \overline{1, n}, \quad l = \overline{1, q};$

- c_{ikl} – вартість перевезення одиниці l -го виду вантажу з i -го пункту відправлення в k -й перевалочний пункт, $i = \overline{1, m}, \quad k = \overline{1, p}, \quad l = \overline{1, q};$

- c_{kjl} – вартість перевезення одиниці l -го виду вантажу з k -го перевалочного пункту в j -й пункт призначення, $k = \overline{1, p}, \quad j = \overline{1, n}, \quad l = \overline{1, q};$

- t_{il} – витрати на виробництво l -го виду вантажу в i -му пункті відправлення;

- b_{jl} – потреба в l -м виді вантажу в j -му пункті призначення;

- c_{kl} – місткість k -го перевалочного пункту стосовно l -го вантажу;

- x_{ijl} – кількість l -го виду вантажу перевезеного з i -го пункту відправлення в j -й пункт призначення;

- y_{ikl} – кількість l -го виду вантажу перевезеного з i -го пункту відправлення в k -й перевалочний пункт;

- z_{kjl} – кількість l -го виду вантажу перевезеного з k -го перевалочного пункту в j -й пункт призначення;

- s_{il} – кількість виробленого l -го виду вантажу в i -му пункті відправлення.

Математичну модель задачі, з урахуванням вище наведених значень, можна подати у вигляді задачі лінійного програмування:

$$y = \sum_{i=1}^m \sum_{j=1}^n \sum_{l=1}^q c_{ijl} * x_{ijl} + \sum_{i=1}^m \sum_{k=1}^p \sum_{l=1}^q c_{ikl} * y_{ikl} + \sum_{k=1}^p \sum_{j=1}^n \sum_{l=1}^q c_{kjl} * z_{kjl} + \sum_{i=1}^m \sum_{l=1}^q t_{il} s_{il} \rightarrow \min_{x_{ijl}, y_{ikl}, z_{kjl}, s_{il} \in \Omega}; \quad (2.50)$$

$$\Omega: \sum_{j=1}^n x_{ijl} + \sum_{k=1}^p y_{ikl} \leq s_{il}, \quad i = \overline{1, m}, \quad l = \overline{1, q}; \quad (2.51)$$

$$\sum_{i=1}^m x_{ijl} + \sum_{k=1}^p z_{kjl} \geq b_{jl}, \quad j = \overline{1, n}, \quad l = \overline{1, q}; \quad (2.52)$$

$$\sum_{i=1}^m y_{ikl} \leq c_{kl}, \quad k = \overline{1, p}, \quad l = \overline{1, q}; \quad (2.53)$$

$$\sum_{i=1}^m y_{ik} = \sum_{j=1}^n z_{kj}, \quad k = \overline{1, p}, \quad l = \overline{1, q}; \quad (2.54)$$

$$x_{ijl} \geq 0; \quad y_{ikl} \geq 0; \quad z_{kjl} \geq 0, \quad s_{il} \geq 0, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad k = \overline{1, p}, \quad l = \overline{1, p}. \quad (2.55)$$

Цільова функція в математичній моделі (2.50) – (2.55) відрізняється від цільової функції (2.44) тільки тим, що враховує витрати на виробництво продукції (вантажу) в пунктах відправлення.

Транспортна задача про закриття підприємства.

Змістовна постановка задачі. Виробниче об'єднання складається з заводів і складів. Задано потреби складів у продукції та вартості на перевезення продуктів з кожного заводу на кожний склад. Задані фіксовані вартості функціонування заводів і можливості заводів по виробництву продукту. Виробниче об'єднання розглядає можливість закриття одного або декількох заводів. Це повинно зменшити витрати на перевезення. Які заводи, якщо це доцільно, повинні бути закриті?

Математичне формулювання завдання. Позначимо:

- c_{ij} – вартості перевезення j -го заводу на i -й склад;
- d_i – потреби i -го складу в продукції, $i = \overline{1, \dots, m}$;
- a_j – можливість j -го заводу по виробництву продукту;
- e_j – фіксована вартість функціонування j -го заводу;

- z_j – булеве число, що показує чи потрібно закрити j -й завод (значення дорівнює 0), чи залишити його працювати (значення дорівнює 1);
- x_{ij} – кількість перевезеного товару з j -го заводу на i -й склад.

Тоді математична модель транспортної задачі про закриття заводу може бути подана у такому вигляді:

$$y = \sum_{j=1}^n \left(z_j * e_j + \sum_{i=1}^m c_{ij} x_{ij} \right) \rightarrow \min_{z_j, x_{ij} \in \Omega}; \quad (2.56)$$

$$\Omega: \sum_{i=1}^m x_{ij} \leq z_j a_j, \quad j = \overline{1, n}; \quad (2.57)$$

$$\sum_{j=1}^n x_{ij} \geq d_i, \quad i = \overline{1, m}; \quad (2.58)$$

$$x_{ij} \geq 0, \quad i = \overline{1, m}; \quad j = \overline{1, n}; \quad (2.59)$$

$$z_j \in \{0, 1\}; \quad j = \overline{1, n}. \quad (2.60)$$

Тут цільова функція (2.56) визначає загальні витрати виробничого об'єднання на функціонування заводів і транспортування готової продукції на склади. Обмеження (2.57) визначає можливості заводів з виробництва продукції. Обмеження (2.58) визначає потреби складів у готовій продукції.

2.2 Теорія достовірності та аналіз лінійних моделей оптимізаційних задач

Для розуміння теоретичних аспектів і прикладних напрямів використання теорії достовірності важливе значення має визначення поняття «достовірність».

Термін *достовірність* використовується в теорії ймовірності, логіці, гносеології та праві, визначається як характеристика обґрунтованого, доказового, безперечного та як синонім істини [28].

Існує ще декілька визначень поняття «достовірність».

Достовірність – це форма існування істини, яка обґрунтована будь-яким способом [28]. В словнику термінів логіки *достовірність* визначається, як обґрунтованість, доказовість знання.

З позиції теорії ймовірності *достовірність* визначається як поняття, що відображає упевненість суб'єкта в правильності своєї оцінки ймовірності настання тієї або іншої події [28].

Теорія достовірності спрямована на вирішення проблеми визначення достовірності проведеного дослідження для подальшого використання його результатів. У цьому аспекті важливе значення має оцінка достовірності інформації, яка використовується в дослідженні, і оцінка достовірності результатів дослідження.

Оцінка достовірності інформації – це складне завдання. Слід зазначити, що, наприклад, експерт, який формулює висновки щодо досліджуваних питань, може помилитись або свідомо вводити в оману. Тому необхідно не тільки оцінювати подану інформацію, а і її джерела. У зв'язку з цим вводяться такі терміни: авторизованість, авторитетність, інформативність та свідомість.

Авторизованість – це прив'язка наданих даних до визначеного джерела. Як джерело можуть бути використані посилання на сайт, книгу та інше [29].

Авторитетність – характеристика джерела даних. Якщо експерт не посилається на джерело, а говорить від першої особи, то оцінюється авторитетність самого експерта [30].

Інформативність показує кількість нових даних, які відносяться до розглянутого питання [31].

Свідомість – показує, наскільки зрозуміло висвітлені питання [31].

При оцінці достовірності результатів дослідження необхідно враховувати те, що достовірним можна вважати результат, допустима погрішність якого не виходить за різницю між розрахунковим значенням моделі і отриманим значенням:

$$|\hat{y} - y_i| \leq \varepsilon, \quad (2.61)$$

де \hat{y} – розрахункове значення моделі;

y_i – вихідне значення показника моделі;

ε – величина допустимої погрішності.

Саме виникнення значної погрішності результатів дослідження призводить до зниження достовірності результатів моделювання.

У цілому погрішність виникає у зв'язку з [32]:

- неадекватністю моделі;
- помилками розрахунку показників і параметрів моделі;
- низькою якістю інформації, яка використовується для моделювання.

Для визначення факту наявності або відсутності помилок, використовують спосіб, який полягає в порівнянні рішення з аналітичними рішеннями. Цей спосіб використовується при наявності значної кількості різних аналітичних рішень. Проте, екстраполювати результати оцінки одного порівняння на всі можливі рішення та параметри моделі досить небезпечно, оскільки в кожному конкретному випадку можуть виникати відхилення, особливо при зміні параметрів [32].

Досить часто використовується спосіб, коли розрахунок одного і того ж параметру відбувається декількома способами. Цей спосіб дає можливість визначити погрішності практично для всіх комбінацій параметрів моделі.

Розглянемо приклад визначення достовірності на основі розробленої математичної моделі. Розроблена лінійна модель виду:

$$\frac{K'}{K} = 0,326 - 0,212 \cdot P_{\text{в}} - 0,08 \cdot P_{\text{з}}, \quad (2.62)$$

де $\frac{K'}{K}$ – індикатор розвитку (відношення нового капіталу до інвестованого капіталу);

$P_{\text{в}}$ – рівень витрат;

$P_{\text{з}}$ – рівень матеріальних запасів.

Встановимо достовірність розрахунків моделі (2.62) на основі встановленої погрішності. Вихідні статистичні дані показників представлено в таблиці 2.6.

Таблиця 2.6 – Вихідні статистичні дані показників моделі (2.62)

№ спостереження	$\frac{K'}{K}$	$P_{\text{в}}$	$P_{\text{з}}$
1	2	3	4
1	0,113	0,957	0,042
2	0,09	-1,052	0,057
3	0,092	1,047	0,052
4	0,083	1,075	0,066
5	0,069	1,113	0,09
6	0,067	1,123	0,099

Продовження таблиці 2.6

1	2	3	4
7	0,069	1,118	0,094
8	0,08	1,085	0,071
9	0,081	1,09	0,066
10	0,009	1,354	0,137
11	0,016	1,33	0,127
12	0,024	1,302	0,118
13	0,031	1,269	0,113
14	0,048	1,198	0,108
15	0,05	1,193	0,104
16	0,055	1,174	0,099
17	0,057	1,165	0,099
18	0,063	1,146	0,094
19	0,063	1,141	0,094
20	0,074	1,113	0,075
21	0,075	1,108	0,071
22	0,08	1,09	0,066
23	0,082	1,08	0,061

Розв'язання

Визначення оцінок індикатора розвитку та помилки розрахунків моделі (2.62) представимо в таблиці 2.7.

Таблиця 2.7 – Розрахунок оцінок індикатора розвитку та помилки

№ спостереження	$\frac{K'}{K}$	Параметри моделі (2.62)				$\sum ((\frac{K'}{K})' - \frac{K'}{K})$
		0,326	$-(0,212 P_B)$	$-(0,08 P_3)$	$(\frac{K'}{K})'$	
1	2	3	4	5	6	7
1	0,113	0,326	-0,203	-0,009	0,114	0,001
2	0,09	0,326	-0,223	-0,012	0,091	0,001
3	0,092	0,326	-0,222	-0,011	0,093	0,001
4	0,083	0,326	-0,228	-0,014	0,084	0,001
5	0,069	0,326	-0,236	-0,019	0,071	0,002
6	0,067	0,326	-0,238	-0,021	0,067	0
7	0,069	0,326	-0,237	-0,020	0,069	0
8	0,08	0,326	-0,230	-0,015	0,081	0,001

Продовження таблиці 2.7

1	2	3	4	5	6	7
9	0,081	0,326	-0,231	-0,014	0,081	0
10	0,009	0,326	-0,287	-0,029	0,01	0,001
11	0,016	0,326	-0,282	-0,027	0,017	0,001
12	0,024	0,326	-0,276	-0,025	0,025	0,001
13	0,031	0,326	-0,269	-0,024	0,033	0,002
14	0,048	0,326	-0,254	-0,023	0,049	0,001
15	0,05	0,326	-0,253	-0,022	0,051	0,001
16	0,055	0,326	-0,249	-0,021	0,056	0,001
17	0,057	0,326	-0,247	-0,021	0,058	0,001
18	0,063	0,326	-0,243	-0,020	0,063	0
19	0,063	0,326	-0,242	-0,020	0,064	0,001
20	0,074	0,326	-0,236	-0,016	0,074	0
21	0,075	0,326	-0,235	-0,015	0,076	0,001
22	0,08	0,326	-0,231	-0,014	0,081	0,001
23	0,082	0,326	-0,229	-0,013	0,084	0,002
Σ	1,471	x	x	x	1,492	0,021

В результаті розрахунків встановлено, що помилка розрахунків складає 0,021 тис. грн. або 2,1 %. Таке значення свідчить про високий рівень достовірності розрахунків параметрів лінійної моделі, оскільки значення помилки наближається до нуля.

З теорії статистики відомо, якщо рівень помилки складає від 0 до 5 %, то результати розрахунків можна вважати достовірними.

Слід вказати, що для визначення ступеня достовірності результатів математичного дослідження необхідно для кожної відносної або середньої величини розрахувати відповідну *середню помилку* [33].

Середня помилка дозволяє визначити межі, в яких з відповідною ймовірністю може знаходитись значення показників розробленої моделі. При оцінці достовірності визначається і середня помилка різниці між двома середніми або відносними величинами [33]:

$$M_{\text{різниці}} = \sqrt{M_1^2 + M_2^2}, \quad (2.63)$$

де M_1^2, M_2^2 – квадрати середніх помилок.

Якщо різниця середніх величин більше середньої помилки різниці в 2,5 – 3,0 рази, то з відповідною ймовірністю можна стверджувати, що різниця

середніх (відносних) величин не випадкова, а залежить від будь-якої визначеної причини.

Для встановлення достовірності розрахунків необхідно оцінити середню помилку розрахунків, а також граничну помилку розрахунків. Це потрібно для виявлення довірчих границь.

Довірчі границі визначаються [34]:

$$\hat{y} - \varepsilon_{zy} \leq y \leq \hat{y} + \varepsilon_{zy}, \quad (2.64)$$

де \hat{y} – розрахункові значення показника лінійної моделі;

ε_{zy} – граничне значення помилки розрахунків;

y – вихідні значення показника.

Для визначення граничної помилки розрахунків використовують таке співвідношення:

$$\varepsilon_{zy} = 1,96 \cdot \bar{\varepsilon}, \quad (2.65)$$

де $\bar{\varepsilon}$ – середня помилка розрахунків.

Середня помилка розрахунків визначається за формулою:

$$\bar{\varepsilon} = \sigma_y \sqrt{1 - \eta_{y/x}^2}, \quad (2.66)$$

де σ_y – середнє квадратичне відхилення показника;

$\eta_{y/x}$ – емпіричне кореляційне відношення.

Для оцінки зв'язку між змінними використовується *емпіричне кореляційне відношення* ($\eta_{y/x}^2$) [35], яке є часткою дисперсії функції y за рахунок впливу аргументу x . У даному випадку загальна (повна) дисперсія розкладається на дві частини – дисперсію усередині кожного інтервалу зміни функції $\sigma_{y/x}^2$, яка не залежить від впливу x , і дисперсію середніх значень функції $\delta \frac{2}{y}$, яка викликана впливом аргументу, тобто

$$\sigma_y^2 = \sigma_{y/x}^2 + \delta \frac{2}{y}. \quad (2.67)$$

Формула для оцінки зв'язку між змінними має вигляд:

$$\eta_{y/x}^2 = \frac{\sigma_{\bar{y}}^2}{\sigma_y^2} = \frac{\frac{1}{n} \sum_{i=1}^n (\tilde{y}_i - \bar{y})^2}{\sigma_y^2}, \quad (2.68)$$

а в разі згрупованих даних:

$$\eta_{y/x}^2 = \frac{\frac{1}{n} \sum_{i=1}^n (\tilde{y}_i - \bar{y})^2 m_i}{\sigma_y^2}, \quad (2.69)$$

де \tilde{y}_i – розрахункове значення функції;

\bar{y} – середнє значення функції за вибіркою;

n – обсяг вибірки;

m_i – кількість спостережень y в кожному інтервалі зміни.

Кореляційне відношення не залежить від одиниць вимірювання змінних, що вивчаються. Воно показує, яку частину загальної дисперсії σ_y^2 можна віднести за рахунок зміни аргументу на одну σ_x^2 . [36].

При цьому характеристика $\eta_{y/x}^2$ тим точніше визначає частку впливу x на загальну дисперсію y , чим менше варіюється залишкова дисперсія $\sigma_{y/x}^2$ при кожному x . Якщо $\eta_{y/x}=1$, то має місце функціональна залежність y від x . Якщо $\eta_{y/x}=0$ – y кореляційно не залежить від x .

Слід зазначити, що результати моделювання також вважаються адекватними або достовірними, якщо обґрунтовано, що вибірки реальних значень показників системи і отриманих результатів мають однакові закони розподілу. Тобто, спільні функції розподілу векторів параметрів, які характеризують умови функціонування системи і моделі, і векторів вихідних характеристик дорівнюють один одному.

Для визначення міри різниці між функціями розподілу вводиться *міра достовірності результатів моделювання* або *міра відмінності моделі або системи* [36].

Розрахунок міри відмінності не складає особливих труднощів в трьох випадках. Перший випадок, коли повністю відомі закони розподілу показників ефективності, по яких порівнюються система і модель, другий – коли відомі

закони розподілу показників системи і моделі до параметрів, третій – якщо є достатній обсяг вибірки результатів моделювання.

Для всіх трьох випадків існують класичні методи оцінки достовірності результатів моделювання. Так, для одновимірного показника оцінка достовірності результатів моделювання може проводитись за допомогою *статистичних критеріїв Уїлкоксона, Пірсона, Колмогорова-Смірнова, метода відношення правдоподібності* [37].

Для оцінки достовірності результатів моделювання по векторному показнику можна скористатися *методом відношення правдоподібності або розпізнавальними системами* визначення ступеня близькості класів, що розпізнають. Як розпізнавальні класи використовуються генеральні сукупності векторів-показників системи $\{\vec{y}\}_c$ і моделі $\{\vec{y}\}_m$, які описуються ймовірностями P_c і P_m [38].

Сутність класичних методів оцінки достовірності зводиться до перевірки гіпотези про розбіжність результатів моделювання і вихідних даних для заданого рівня значущості, на якому перевіряється гіпотеза.

Розглянуті методи оцінки достовірності припускають наявність вибірок (статистики) результатів моделі і системи, оскільки по ним можна визначити оцінки функцій розподілу показників моделі і системи.

Слід вказати і на те, що результати математичного моделювання подібні результатам реальних подій на рівні випадкових ймовірних подій, тому можна стверджувати, що кількісні результати моделювання подій в цілому є достовірними.

Для дослідження достовірності результатів моделювання використовують *методи сценаріїв* за вибірками [39]. Причому сценарій складається таким чином, щоб умови здійснення кожної ймовірнісної події в моделі співпадали з умовами, в яких проводились дослідження.

Для оцінки достовірності результатів моделювання за малою вибіркою може бути використана методика перевірки подібності теоретичної і емпіричної функцій розподілу, яка заснована на використанні *модифікованих критеріїв Колмогорова-Смірнова і Мізеса* [40]. Проте ці методи використовуються для незначного класу задач.

Існують і інші підходи для вирішення задачі оцінки достовірності по малій вибірці, відповідно до яких за наявності малої вибірки доцільно говорити не про оцінку достовірності результатів моделювання, а про оцінку статистичної несуперечності результатів дослідження результатам моделювання.

При цьому під *статистичною несуперечністю* слід розуміти несуперечність результатів дослідження твердженню, що їх значення на заданому рівні значущості можуть розглядатися як ті, що належать до тієї ж генеральної сукупності, що і вибірка результатів моделювання [30].

Якщо реалізація досліджених процесів не відповідає закону розподілу результатів моделювання (змінюють його параметри), то вони належать до іншої генеральної сукупності.

Для перевірки достовірності результатів моделювання процесів можна використати *перевірку на статистичну еквівалентність моделей елементарних подій*. Цей метод дозволяє оцінити достовірність результатів елементарних подій моделі за статистикою, яка отримана за спеціальними моделями елементарних подій [31].

Таким чином, представлені методики оцінки достовірності інформації та результатів дослідження є підґрунтям для отримання моделі, яка віддзеркалює реальні події і може бути використана для прийняття ефективних управлінських рішень.

Аналіз лінійних моделей оптимізаційних задач спрямований на прийняття оптимального рішення.

Лінійна оптимізаційна модель включає [32]:

- систему обмежень;
- цільову функцію;
- області допустимих рішень;
- критерії оптимальності.

Цільова функція в загальному вигляді складається з трьох елементів [23]:

- змінних, які управляються;
- змінних, які не управляються;
- форми функції (виду залежності між змінними).

Область допустимих рішень – це область, в межах якої здійснюється вибір рішень. У задачах вона обмежена наявними ресурсами і умовами, які записуються у вигляді системи обмежень, які складаються із рівнянь [19].

Критерії оптимальності – це показник, який визначається за допомогою цільової функції через інші показники [23]. Одному і тому ж критерію оптимальності можуть відповідати декілька різних, але еквівалентних цільових функцій. Моделі з однією і тією ж системою обмежень можуть мати різні критерії оптимальності й різні цільові функції.

Для здійснення аналізу лінійних моделей необхідно побудувати математичну модель, методика розробки якої полягає в тому, щоб сутність

задачі представити математично, використовуючи різні символи, змінні й постійні величини, індекси та інші символи.

Всі умови задачі записують у вигляді рівнянь. Тому, в першу чергу, необхідно визначити систему змінних, які можуть для конкретної задачі висвітлювати вихідні значення показників.

Слід зазначити, що будь-яку оптимізаційну задачу лінійного програмування можна привести до задачі лінійного програмування в канонічній формі. Для цього в загальному випадку необхідно зводити задачу максимізації до задачі мінімізації, переходити від обмежень нерівностей до обмежень рівнянь і замінювати змінні, які не підходять умовам.

Правило приведення оптимізаційної задачі лінійного програмування до канонічного вигляду полягає в такому [23 – 26]:

- якщо в початковій задачі вимагається визначити максимум лінійної функції, то слід змінити знак і шукати мінімум цієї функції;
- якщо в обмеженнях права частина «-», то слід помножити це обмеження на -1;
- якщо серед обмежень є нерівності, то шляхом введення додаткових від'ємних змінних вони перетворюються в рівняння;
- якщо деяка змінна x_i не має обмежень по знаку, то вона замінюється (в цільовій функції і у всіх обмеженнях) різницею між двома новими від'ємними змінними.

Складність вирішення оптимізаційних задач лінійного програмування, побудови відповідних моделей та їх аналізу залежить [31]:

- від виду функціональних залежностей, тобто від зв'язку функції з елементами розв'язку;
- від розмірності задачі, тобто від кількості елементів розв'язку;
- від виду і кількості обмежень, які накладаються на елементи розв'язку.

Під час аналізу лінійних оптимізаційних моделей важливим етапом є *інтерпретація* отриманих результатів. Саме на цьому етапі проявляється кваліфікація спеціалісту з оцінки землі та нерухомого майна. Інтерпретація полягає в тому, що на базі розробленої лінійної оптимізаційної моделі визначаються зв'язки між показниками, оптимізаційні розв'язки проблем, пропонуються управлінські рішення.

2.3 Цілочислове програмування

Цілочислове програмування – це різновид задач лінійного програмування, в якому змінні та отримані результати повинні бути цілими числами [27].

Задачі цілочислового програмування можуть бути лінійними, якщо обмеження і цільова функція задачі є лінійною залежністю, або нелінійними – якщо залежності будуть мати нелінійну форму.

Слід вказати, що широке використання задач цілочислового програмування полягає в тому, що необхідно отримувати цілочислове рішення.

Цілочислове програмування виникло в 50 – 60-ті роки 20 століття на основі робіт Дж. Данцига і Р. Гоморі [25].

Задача цілочислового програмування записується так [26]:

$$Z = \sum_{j=1}^n c_j x_j \rightarrow \max(\min), \quad (2.70)$$

за умови того, що x_j є цілим додатним числом.

Для знаходження оптимального вирішення цілочислових задач застосовують відповідні методи. Найпростішим методом вирішення цілочислової задачі є знаходження її оптимального розв'язку як задачі, що має лише неперервні змінні, з подальшим округленням останніх.

Для знаходження оптимальних планів задач цілочислової програмування застосовують дві основні групи методів [33]:

- методи відтинання;
- комбінаторні методи.

Основою методів відтинання є ідея поступового «звуження» області допустимих шляхів вирішення задачі. Пошук цілочислового оптимального вирішення задачі починається з розв'язання задачі з так званими послабленими обмеженнями, тобто без урахування вимог цілочисленності змінних. Далі вводять у модель спеціальні додаткові обмеження, що враховують цілочисленність змінних, багатокутник допустимих рішень послабленої задачі поступово зменшується до отримання змінних оптимального вирішення (до отримання цілочислових значень).

До цієї групи належать [28]:

- методи розв'язання повністю цілочислових задач (дробовий алгоритм Гоморі);
- методи розв'язання частково цілочислових задач (другий алгоритм Гоморі, або змішаний алгоритм цілочислового програмування).

Комбінаторні методи цілочислової оптимізації базуються на повному переборі всіх допустимих цілочислових рішень, тобто вони реалізують процедуру цілеспрямованого перебору, під час якої розглядається лише частина розв'язків (досить невелика), а решта враховується одним зі спеціальних методів. Найпоширенішим у цій групі методів є метод віток і меж.

Рекомендації по формулюванню і вирішенню задач цілочислового програмування [32]:

- кількість цілочислових змінних зменшувати наскільки можливо. Наприклад, цілочислові змінні, значення яких повинно бути не менше 20, можна розглядати як безперервні;

- на відміну від загальних задач лінійного програмування, додавання нових обмежень, особливо включених цілочислових змінних, звичайно зменшує час вирішення задач цілочислового програмування;

- немає необхідності в знаходженні точного оптимального цілочислового рішення, відмінного від безперервного рішення, наприклад, на 3 %.

Тоді реалізацію методу гілок і меж для задачі максимізації можна закінчувати, якщо відношення різниці між верхньою і нижньою межами до верхньої межі менше 0,03.

Алгоритм розв'язання задач цілочислового програмування такий [34]:

- розв'язується задача лінійного програмування без обмежень на цілочисельність, наприклад, симплекс-методом;

- якщо оптимальне рішення задачі лінійного програмування нецілочисельне, то проводиться «велика ітерація». Будується лінійне обмеження, якому задовольняє будь-яке цілочислове рішення задачі і не задовольняє отримане оптимальне нецілочисельне значення.

Геометрично це означає – провести перетин (гіперплощини), який відсікав би нецілочислову вершину, не зачіпаючи решту цілочислових крапок. Такий перетин називають *правильним*. Правильний перетин повинен задовольняти таким умовам [39]:

- умова відсікання: оптимальне рішення задачі лінійного програмування не задовольняє умові відсікання;

- умова правильності: всі цілочислові рішення задачі задовольняють умові відсікання.

Оскільки для початкової задачі додаткове обмеження (відсікання) даватиме неприпустиме базисне рішення, необхідно провести «малі ітерації» для отримання оптимального рішення. Якщо отримане оптимальне рішення

нецілочисельне, то проводиться наступне відсікання. В іншому випадку пошук завершений.

Р. Гоморі запропонував ідею формування додаткових обмежень, яка призводить до вирішення задач цілочислового програмування за кінцеве число кроків [40].

Метод Гоморі.

Перший алгоритм Р. Гоморі.

Нехай задана повністю цілочислова лінійна задача:

$$\min z(\bar{x}) = \sum_{j=1}^n c_j x_j, \quad j = 1, \dots, n, \quad (2.71)$$

$$\begin{cases} \sum_{i=1}^m a_{ij} x_j = b_i, & i = 1, \dots, m, \\ x_j \geq 0, \\ x_j \in \mathbb{Z}. \end{cases} \quad (2.72)$$

Нехай в результаті лінійних операцій над рівняннями системи (2.72) отримано нове лінійне рівняння:

$$\sum_{j=1}^n a_j x_j = b. \quad (2.73)$$

Цілою частиною числа a називається найбільше ціле число $[a]$, яке не перевищує a . Наприклад $[4,25] = 4$; $[-4,25] = -5$.

Замінивши в рівнянні (2.73) всі коефіцієнти a_j їх цілими частинами, отримаємо:

$$\sum_{j=1}^n [a_j] x_j \leq b. \quad (2.74)$$

Якщо всі x_k є цілочисловими, то ліва частина рівняння (2.74) теж цілочислова. Рівняння (2.74) можна посилити таким чином:

$$\sum_{j=1}^n [a_j] x_k \leq [b]. \quad (2.75)$$

Це обмеження-нерівність (2.75) можна перетворити в обмеження-рівняння шляхом введення додаткової цілочислової і від'ємної змінної x :

$$\sum_{j=1}^n [a_j] x_j + x = [b]. \quad (2.76)$$

Віднімемо (2.73) з (2.76). Враховуючи, що $a = [a] + \{a\}$ отримаємо:

$$x = -\{b\} - \left(\sum_{j=1}^n -\{a_j\} x_j \right). \quad (2.77)$$

У першому методі Гоморі всі обмеження формуються відповідно до рівнянь (2.77) – це і є відсікання нецілочислових крапок.

Фактично, (2.77) – рівняння *відсікаючої гіперплощини*. Після вирішення задачі лінійного програмування одержуємо деяке базисне рішення, в якому деякі змінні можуть бути цілочисловими, а інші – нецілочисловими [38].

Для нецілочислових базисних змінних будується відсікання по першому алгоритму Гоморі, послідовно, для кожної з них. Якщо нецілочислових змінних декілька, то краще вибрати ту, у якої більша дробова частина.

Якщо рівняння вибраної базисної змінної має вигляд:

$$x_{\text{баз}} = b_i - (\sum a_{ij} x_j), \quad (2.78)$$

то рівняння додаткового обмеження-відсікання виглядає таким чином:

$$x = -\{b_i\} - \left(\sum -\{a_{ij}\} x_j \right). \quad (2.79)$$

Побудова відсікання – «велика ітерація», далі для звуженої області проводяться «малі ітерації».

Вирішення задач цілочислового лінійного програмування може бути відсутнє, якщо:

- цільова функція є необмеженою знизу, тобто як початкова нецілочислова, так і цілочислова задача лінійного програмування не мають рішення;
- початкова задача лінійного програмування має рішення, а цілочислова не має. В симплекс-таблиці це видно таким чином: в стовпці вільних членів в деякому рядку стоїть неціле число, а вся решта коефіцієнтів цього рядка – цілі числа.

Збіжність даного алгоритму достатньо повільна. Для задачі з 10-ма змінними необхідно провести біля 1000 ітерацій.

Перебір можливих цілих точок допустимої області не зменшує обчислень, оскільки їх звичайно достатньо багато.

Округлення нецілочислового рішення до найближчого цілого може дати крапку, що не належить області.

Другий алгоритм Р. Гоморі.

Застосовується для вирішення як повністю, так і не повністю цілочислових задач. Багато індексів I при змінних x_j розбиваються на дві неперехресні підмножини I_1 і I_2 так, що:

при $j \in I_1, x_j \in Z$ (x_j – цілочислові);

при $j \in I_2, x_j \in R$ (x_j – нецілочислові).

Як і в першому алгоритмі Гоморі, спочатку знаходиться вирішення задачі лінійного програмування без обмежень на цілочисельність, і далі проводяться відсікання за допомогою введення додаткових обмежень.

Припустимо, що після розв'язання задачі лінійного програмування отримана деяка базисна змінна, яка повинна бути цілочисловою:

$$x_5 = b_i - (\sum_j a_{ij} x_j), \quad (2.80)$$

де x_j – вільна змінна.

За другим алгоритмом Гоморі відсікання будується таким чином:

$$x_5 = -\{b_i\} - (\sum_j \alpha_{ij} x_j), \quad (2.81)$$

де

$$\alpha_{ij} = \begin{cases} a_{ij}, & a_{ij} \geq 0 \\ \frac{\{b_i\}}{1 - \{b_i\}} |a_{ij}|, & a_{ij} < 0 \end{cases} \left\{ x_j \in R; \right. \quad (2.82)$$

$$\left. \begin{cases} \{a_{ij}\}, & \{a_{ij}\} \leq \{b_i\} \\ \frac{\{b_i\}}{1 - \{b_i\}} (1 - \{a_{ij}\}), & \{a_{ij}\} > \{b_i\} \end{cases} \right\} x_j \in Z.$$

Метод віток і меж.

Метод віток і меж застосовується як до повністю цілочислових задач, так і до частково цілочислових задач [34].

Спочатку розв'язується ослаблена задача без обмежень на цілочисельність x_k повинна бути цілочисловою:

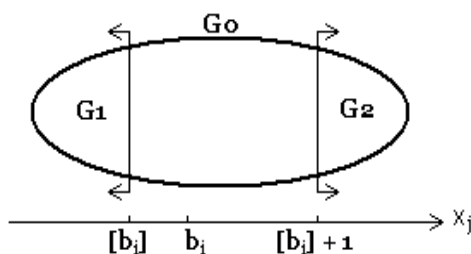
$$x_k = b_k - \left(\sum_{j=1}^n a_{kj} x_j \right). \quad (2.83)$$

Тоді радіус допустимих рішень G_0 розбивається на дві підмножини G_1 і G_2 таким чином, що:

$$G_1: x_j \leq [b_i],$$

$$G_2: x_j \geq [b_i] + 1. \quad (2.84)$$

Таким чином, виключається інтервал $[b_i] + 1 > x_j > [b_i]$, який не містить цілочислових точок.



Початкова задача розбивається на дві підзадачі. В кожній з областей G_1 , G_2 знаходяться оптимальні точки, і, якщо вони не задовольняють умові цілочисловості, знову здійснюється розгалуження. Якщо отриманий оптимум виявляється допустимим

(цілочисельним) для даної задачі, він фіксується і оголошується як найкращий. При цьому немає необхідності продовжувати розгалуження цієї підзадачі, оскільки поліпшити це рішення не вдасться.

Як тільки отримане допустиме (цілочислове) рішення іншої підзадачі виявляється кращим, воно фіксується замість попереднього.

Процес розгалуження продовжується до тих пір, поки кожна підзадача не призведе до цілочислового рішення, або не буде встановлена неможливість поліпшення наявного рішення. Висновок про необхідність подальшого розбиття задачі робиться на основі введення межі.

Як межа використовується значення цільової функції отриманого допустимого цілочислового рішення. Якщо будь-яке оптимальне рішення

підзадачі забезпечує гірше значення цільової функції, ніж наявне рішення (прийняте як межа), то цю підзадачу розглядати далі не слід.

Алгоритм методу віток і меж є ітеративним.

t – номер ітерації; z_t – оцінка (для задачі мінімізації $z_0 = \infty$).

При використанні цього алгоритму потрібне вирішення послідовності задач лінійного програмування без обмежень на цілочисельність.

Послідовність задач, що підлягають вирішенню називається основним списком [19]:

1. Якщо основний список порожній – закінчення алгоритму, інакше – вибирають задачу з основного списку і знаходять її оптимальне рішення.

2. Якщо вибрана задача не має рішення, або її оптимальне рішення гірше прийнятої оцінки, то необхідно виключити цю задачу із списку і перейти до п. 1, інакше – до п. 3.

3. Якщо отримане рішення цілочислове – сформувати нову оцінку z_{t+1} , яка відповідає якнайкращому оптимальному рішенню поточної задачі. Перехід до п. 1. Якщо рішення поточної задачі нецілочислове, то оцінка залишається у такому вигляді: $z_{t+1} = z_t$.

4. Вибирається одна із змінних x_k , яка за умовою повинна бути цілочисловою. Проводиться розгалуження, тобто в основний список додається дві підзадачі, для яких зберігаються ті ж обмеження, але для однієї:

$$x_k \leq [b_k], \quad (2.85)$$

а для іншої:

$$x_k \geq [b_k] + 1. \quad (2.86)$$

Перехід до п. 1.

Слід зазначити, що вибір змінної може бути довільним (за збільшенням номерів) або визначатися таким чином:

- представлена змінна є важливим рішенням, що приймається в рамках розробленої моделі;
- коефіцієнт в цільовій функції істотно перевершує всі інші.

Цілочислова транспортна задача.

В попередньому розділі розглядалися математичні аспекти і особливості вирішення транспортних задач. Одним з видів цих задач є цілочислова транспортна задача, яка має цілочисловий характер. *Цілочислові змінні мають місце*, коли перевезений вантаж є лічильною множиною великих заготівель або комплектуючих, неподільних продуктів виробництва, упакованих сипучих матеріалів і т.п. Об'єм такого вантажу характеризується розміром, що виражається в штуках, пакунках, партіях і т.п.

Тоді математична постановка транспортної задачі планування перевезень набуває вигляду:

$$y = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min, \quad x_{ij} \in \Omega \quad (2.87)$$

$$\Omega: f_j = \sum_{i=1}^m x_{ij} = b_j, \quad i=1, m, j=1, n, \quad (2.88)$$

$$f_{n+i} = \sum_{j=1}^n x_{ij} = a_i, \quad i=1, m, j=1, n, \quad (2.89)$$

$$x_{ij} \geq 0, \quad i=1, m, j=1, n, \quad (2.90)$$

$$x_{ij} = \text{int}, \quad i=1, m, j=1, n. \quad (2.91)$$

Математична модель цілочислової транспортної задачі (2.87) – (2.91) відрізняється від раніше розглянутих математичних моделей транспортної задачі додатковим обмеженням на цілочисельність невідомих x_{ij} (2.91).

Це потребує накладення обмеження цілочисельності на функції

$$f_1, f_2, \dots, f_{n+m}.$$

Необхідно зауважити, що в загальному випадку умова цілочисельності може накладатися і на значення функції цілі y .

Задача цілочислового лінійного програмування.

Розглянемо ще декілька задач, математична модель яких відповідає цілочисловій задачі лінійного програмування.

Оптимальне вирішення нижче приведених задач можна одержати методом відсікання, введенням в задачу додаткових обмежень у вигляді нерівностей.

Задача про розподіл вантажного флоту.

Змістовна постановка завдання. Нехай вантажний флот має у своєму складі судна n типів. Кількість суден типу j дорівнює q_j , а витрати при використанні одного судна типу j у планованому періоді складає $c_j, j=1,2,\dots,n$. Кожне судно має вантажні ємкості m типів (трюми, палуби, танки і т.п.). Вантажопідйомність ємкості i на судні типу j дорівнює $d_{ij}, j=1,2,\dots,m$. Перевезенню підлягають p видів вантажу. Вантаж виду k є в кількості $a_k, k=1,2,\dots,p$. Потрібно вибрати найбільш економічний комплекс транспортних засобів для перевезення вантажу.

Математична модель задачі.

Позначимо:

- x_j – кількість суден j -го типу, $j=1,2,\dots,n$;
- z_{ik} – кількість вантажу виду k , що підлягає завантаженню в ємкість $i, k=1,2,\dots,p$.

Тоді математична модель задачі про розподіл вантажного флоту має вигляд:

$$y = \sum_{j=1}^n c_j x_j \rightarrow \min_{x_j \in \Omega}; \quad (2.92)$$

$$\Omega: \sum_{j=1}^n d_{ij} x_j - \sum_{k=1}^p z_{ik} \geq 0, \quad i = \overline{1, m}; \quad (2.93)$$

$$\sum_{i=1}^m z_{ik} = a_k, \quad k = \overline{1, p}; \quad (2.94)$$

$$0 \leq x_j \leq q_j; \quad j = \overline{1, n}; \quad (2.95)$$

$$x_j = \text{int}, \quad j = \overline{1, n}; \quad (2.96)$$

$$z_{ik} \geq 0; \quad i = \overline{1, m}; \quad k = \overline{1, p}. \quad (2.97)$$

Тут обмеження (2.93) показує, що загальна кількість вантажу, яка завантажена в ємкості кожного типу, не повинна перевищувати сумарну вантажопідйомність цих ємкостей у всіх судах, а обмеження (2.94) говорить про те, що перевезення усіх вантажів повинно бути повністю здійснено.

Задача про розвезення вантажу.

Змістова постановка задачі. Нехай деяка центральна база постачає продукцію (її можна вважати однорідною) на m складів. Розвезення продукції на склади здійснюються однією вантажівкою, причому кожний склад одержує своє замовлення в один прийом – вантажопідйомність вантажівки для цього достатня. Вантажівка може одночасно взяти вантаж, що відповідає не більше ніж k замовленням. Вантажівка може об'їжджати склади за визначеними маршрутами. Один і той самий склад може знаходитися на різних маршрутах.

Нехай для кожного складу відома функція витрат, що залежна, наприклад, від розміру замовлення. Потрібно скласти графік розвезень, що забезпечує всіх клієнтів і мінімізує сумарні витрати. Час доставки не враховується. Передбачається, що всі операції по доставці гарантовано можуть бути здійснені протягом деякого періоду часу, що влаштовує всіх споживачів.

Під способом розвезення будемо розуміти будь-яку припустиму комбінацію виконання замовлень, це m -мірний стовпець, 1-й компонент якого дорівнює одиниці, якщо i -те замовлення в цьому способі задовольняється, і нулю – у противному разі. Для будь-якої реальної задачі при невеликих значеннях m можна фактично виписати всі такі способи розвезення. Число n цих способів буде залежати не тільки від перерахованих параметрів, але і від числа складів на кожному маршруті, об'єму замовлень і т.д. Кожному j -му способу розвезення відповідають витрати c_j .

Нехай при даних конкретних умовах задачі сформована матриць $A = [a_{ij}]$ способів розвезення, що складається з нулів і одиниць. Стовпці цієї матриці – способи розвезення, тобто $a_{ij} = 1$, якщо в j -му способі i -те замовлення задовольняється, і $a_{ij} = 0$ – у противному разі. Завдання полягає у виборі найбільш економічної комбінації цих способів.

Математична модель задачі. Введемо змінні: x_j рівні 1, якщо j -й спосіб розвезення реалізується, і рівні 0 – у противному разі. Тоді математична модель задачі набуває вигляду:

$$y = \sum_{j=1}^n c_j x_j \rightarrow \min_{x_j \in \Omega}; \quad (2.98)$$

$$\Omega: \sum_{j=1}^n a_{ij} x_j = 1, \quad i = \overline{1, m}; \quad (2.99)$$

$$x_j \in \{0, 1\}, \quad j = \overline{1, n}; \quad (2.100)$$

$$a_{ij} \in \{0, 1\}, \quad i = \overline{1, m}. \quad (2.101)$$

Умова (2.99) означає, що всі замовлення повинні бути задоволені тільки один раз.

2.4 Нелінійні оптимізаційні моделі в оцінці нерухомого майна

Сучасна наука має в розпорядженні широкий арсенал методів, серед яких особливе значення має математичне моделювання. До останнього часу в дослідженнях прикладного характеру і в управлінській діяльності використовувались лінійні оптимізаційні моделі. При достатньо високому рівні технологічної культури і надійній інформаційній базі реалізація лінійних математичних моделей забезпечує більш високу ефективність в порівнянні з традиційними методами прийняття управлінських рішень. Проте дослідження останніх років показали, що існує клас задач, де застосування лінійних оптимізаційних моделей або некоректно, або неможливо, але в той же час вони ефективно розв'язуються на основі нелінійних математичних моделей.

При визначенні нелінійних зв'язків важливе значення має категорія системи. *Система* (від грецького – складене з частин, з'єднання) – це відносно відособлена і впорядкована сукупність взаємодіючих елементів, здатних реалізувати певні функції [23].

Важливою якістю будь-якої системи є *емерджентність* – наявність таких властивостей, які не властиві жодному з елементів, що входять в систему [33].

Системи бувають детерміновані і стохастичні. Під *детермінованою системою* розуміється система, в якій зв'язки між елементами і подіями строго і однозначно визначені, тобто її стан в майбутньому однозначно визначається станом на відповідний час [39].

Системи, в яких зв'язки між елементами і подіями носять ймовірнісний характер називаються *стохастичними* [39].

Слід зазначити, що більшість процесів мають ймовірнісний характер і можуть бути описані нелінійними моделями. В цілому, введення в детерміновані процеси незначної нелінійності призводить до непередбачуваного ряду подій.

При дослідженні нелінійних зв'язків необхідно враховувати динаміку системи. *Динамічна система* – система, в якій перехід з одного стану в інший здійснюється не миттєво, а протягом деякого часу, тобто процес переходу можна спостерігати і описати [21].

Найбільший інтерес з погляду управління представляють закономірності поведінки складних динамічних систем. Системи, які мають розгалужену структуру і велику кількість взаємозв'язаних і взаємодіючих елементів, що забезпечують виконання будь-якої складної функції, називаються *складними*. Одна з важливих властивостей складних систем – *нелінійність* [18].

Якщо розглядаємо нестійкий режим системи, то, порушивши його, отримаємо деяке відхилення від початкового положення. Відхилення наростатиме до тих пір, поки не вступить в дію механізм нелінійного обмеження даного процесу. З фізичної точки зору це можна пояснити так: наростання амплітуди не може відбуватися до безкінечності, через обмеженість енергетичних ресурсів системи це наростання повинно припинитися або змінитися зменшенням амплітуди відхилення. Будь-який новий режим повинен мати кінцеву амплітуду, і управляють цими процесами нелінійні закони. Властивості нелінійної системи безпосередньо залежать від її стану.

При проведенні аналізу нелінійних динамічних систем необхідно враховувати поняття «фазовий простір» і «траєкторія». Багато станів динамічної системи називають *фазовим простором*, а траєкторію руху в цьому просторі з деякого початкового стану – *фазовою траєкторією*. Найважливіша характеристика цього простору його розмірність – число величин, які необхідно задати для визначення стану системи [40].

Траєкторія – крива в просторі параметрів, яку описує точка при своєму русі в часі [22].

Найважливіше поняття в теорії нелінійних динамічних систем – це поняття *грубості* (або структурної стійкості). Якщо при малих змінах параметрів системи вид фазових змін нелінійної системи залишається незмінним, то таку системи називають *грубою* [24].

При аналізі динаміки нелінійних моделей враховують поняття «*аттрактор*» – множина точок або підпростір у фазовому просторі, до якого наближається траєкторія після загасання перехідних процесів. Класичними

прикладом аттракторів в динаміці можуть служити точки динамічної рівноваги, нерухомі точки відображень, або граничні цикли. Динамічні системи, які володіють аттракторами називають *дисипативними* [27].

При моделюванні нелінійних процесів динаміки необхідно враховувати теорії прогнозу, катастроф і біфуркації. *Теорія катастроф* полягає в відстеженні змін на основі розуміння коли, чому і як відбуваються зміни [31]. Для цих процесів характерні раптові стрибкоподібні зміни, а не поступовий плавний розвиток.

Основи теорії біфуркацій динамічних систем були закладені в працях великого французького вченого Анрі Пуанкаре. *Біфуркація* – зміна характеру руху динамічної системи на великому тимчасовому інтервалі при зміні одного або декількох параметрів [34]. Ті значення параметрів, при яких змінюються якісні або топологічні властивості руху, називаються *критичними або біфуркаційними значеннями* [34]. Хаос в динаміці означає чутливість динамічної еволюції до змін початкових умов.

При вивченні нелінійних об'єктів необхідно враховувати поняття *фракталу* [37]. Будь-який нелінійний процес призводить до розгалуження, в якому система може вибрати ту або іншу гілку. Будь-яка найменша неточність в початкових умовах може пізніше дуже сильно вплинути на подальший рух.

У кожний окремий момент причинний зв'язок зберігається, але після декількох розгалужень його вже не видно.

Поняття *фракталу* розроблялося Бенуа Б. Мандельбротом як альтернатива евклідовій геометрії, що претендувала на найбільш відповідний опис об'єктів природи [37]. Його істотною межею була невичерпність найдрібніших деталей в різних масштабах вимірювання.

Для розробки нелінійних оптимізаційних моделей досліджуваних систем вирішуються задачі нелінійного програмування.

Нелінійне програмування – математичні методи визначення максимуму або мінімуму функції за наявності обмежень у вигляді нерівностей або рівнянь. Максимізувавши або мінімізувавши функцію, визначають критерій ефективності вирішення задачі, цей критерій має назву *цільової функції* [19].

Цільова функція задач нелінійного програмування полягає в тому, щоб знайти умови, що перетворюють цільову функцію в мінімум або максимум. Рішення, що задовольняє умові задачі і відповідає поставленій меті, називається *оптимальним планом* [33]. Нелінійне програмування служить для вибору найкращого плану розподілу обмежених ресурсів в цілях вирішення поставленої задачі. У загальному вигляді постановка задачі нелінійного

У загальному вигляді математична модель задачі нелінійного програмування формулюється таким чином:

$$f = (x_1, x_2, \dots, x_n) \rightarrow \min (\max). \quad (2.102)$$

При цьому ці змінні повинні задовольняти обмеженням:

[illegible]

де одна із функцій f, g_i нелінійна.

Для задач нелінійного програмування немає єдиного методу вирішення. Залежно від виду цільової функції і системи обмежень розроблені спеціальні методи вирішення, до яких відносяться метод множників Лагранжа, градієнтні методи, наближені методи вирішення, графічний метод.

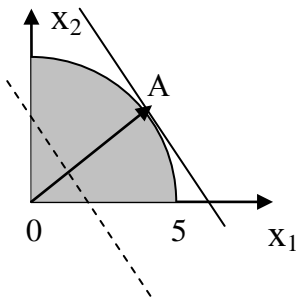
Розглянемо деякі з них. Основні ідеї графічного методу: максимум і мінімум досягається в точках дотику лінії рівня з областю допустимих рішень, яка задається системою обмежень. Наприклад, якщо лінії рівня – прямі, то точки дотику можна визначити, використовуючи геометричне значення похідної.

Розглянемо на прикладах вирішення задач нелінійного програмування.

1. Знайти екстремуми функції $L(x_1, x_2) = x_1 + 2x_2$ при обмеженнях

$$x_1^2 + x_2^2 \leq 25, \ x_1, x_2 \geq 0.$$

Розв'язання



Область допустимого вирішення – це частина кола з радіусом 5, яка розташована в I чверті. Знайдемо лінії рівня функції $L: x_1 + 2x_2 = C$. Виразимо $x_2 = \frac{C}{2} - \frac{x_1}{2}$. Лініями рівня будуть паралельні прямі з кутовим коефіцієнтом, який дорівнює $-\frac{1}{2}$. Мінімум функції

досягається в точці зі значенням $(0;0)$, $L_{min}=0$, оскільки градієнт $\bar{g}(1,2)$ спрямовано вгору вправо. Максимум досягається в точці дотику кривої $x_2 = \sqrt{25 - x_1^2}$ та лінії рівня. Оскільки кутовий коефіцієнт дотику до графіку функції дорівнює $-\frac{1}{2}$, знайдемо координати точки дотику, використовуючи геометричне значення похідної.

$$x_2'(x_0) = -\frac{1}{2}; (\sqrt{25 - x_1^2})' = -\frac{1}{2};$$

$$\frac{-2x_0}{2\sqrt{25 - x_0^2}} = -\frac{1}{2}; \Rightarrow x_0 = \sqrt{5}; x_2 = 2\sqrt{5}.$$

Тоді

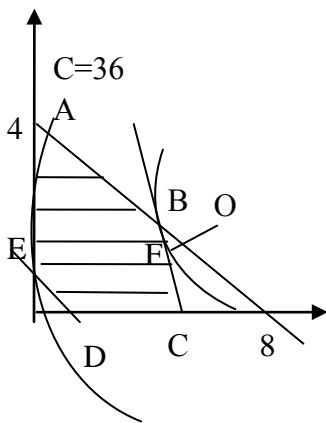
$$L = \sqrt{5} + 2 \cdot 2\sqrt{5} = 5\sqrt{5}.$$

Відповідь: мінімум досягається в точці $O(0;0)$, глобальний максимум дорівнює $5\sqrt{5}$, в точці $A(\sqrt{5}; 2\sqrt{5})$.

2. Знайти екстремуми функції $L = (x_1 - 6)^2 + (x_2 - 2)^2$ при обмеженні:

$$\begin{cases} x_1 + x_2 \leq 8, \\ 3x_1 + x_2 \leq 15, \\ x_1 + x_2 \geq 1, \\ x_1, x_2 \geq 0. \end{cases}$$

Розв'язання



Область допустимого вирішення – багатокутник ABCDE. Лінії рівня представляють собою окружність $(x_1-6)^2+(x_2-2)^2=C$ з центром в точці $O_1(6;2)$. Візьмемо, наприклад, $C = 36$, бачимо, що максимум досягає в точці $A(0;4)$, яка лежить на окружності найбільшого радіусу, який пересікається з областю допустимого вирішення $L(A)=(0-6)^2+(4-2)^2=40$. Мінімум – в точці F , яка знаходиться на перетині прямої $3x_1+x_2=15$ і перпендикуляру до цієї прямої, виведеного із точки O_1 . Оскільки кутовий коефіцієнт дорівнює -3 , то

кутовий коефіцієнт перпендикуляру дорівнює $\frac{1}{3}$. Із рівняння прямої, яка проходить через точку O_1 з кутовим коефіцієнтом $\frac{1}{3}$, отримаємо $(x_2-2)=\frac{1}{3}(x_1-6)$.

Знайдемо координати точки E :

$$\begin{cases} x_1 - 3x_2 = 0, \\ 3x_1 + x_2 = 15. \end{cases}$$

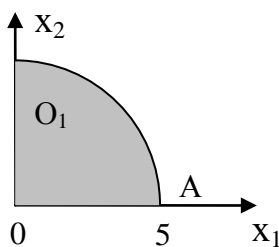
Вирішивши систему, отримаємо точку $E(4.5; 1.5)$.

$$L(E) = (4.5-6)^2 + (1.5-2)^2 = 2.5.$$

Відповідь: мінімум дорівнює 2.5 та досягається в точці $(4.5; 1.5)$, максимум дорівнює 40 в точці $(0;4)$.

3. Знайти екстремуми функції $L=(x_1-1)^2+(x_2-3)^2$ при обмеженнях $x_1^2 + x_2^2 \leq 25$, $x_1, x_2 \geq 0$.

Розв'язання



Область допустимого вирішення – частина кола з центром на початку координат з радіусом 5, яка розташована в I чверті. Лінії рівня – це окружності з центром в точці O_1 і радіусом C , оскільки $(x_1-1)^2+(x_2-3)^2 = C$. Точка O_1 – це розроблена лінія рівня, яка відповідає мініимальному значенню $C = 0$. Глобальний максимум досягається в точці A , яка знаходиться на перетині області допустимого вирішення з лінією рівня найбільшого радіусу.

При цьому $L(A)=(5-1)^2+(0-3)^2=25$.

Відповідь: мінімум дорівнює 0, досягається в точці $(1;3)$, максимум, дорівнює 25, досягається в точці $A(5;0)$.

Сутність методу Лагранжа в побудові функції:

$$L(x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n) + \sum_{i=1}^m \lambda_i g_i(x_1, x_2, \dots, x_n),$$

де λ_i – невідомі для знаходження екстремуму функції L .

Має сенс наступна теорія: якщо точка $(x_1^0, x_2^0, \dots, x_n^0)$ є точкою умовного екстремуму функції $f(x_1, x_2, \dots, x_n)$ за умови $g(x_1, x_2, \dots, x_n) = 0$, то існують значення λ_i такі, що точка $(x_1^0, x_2^0, x_n^0, \lambda_1^0, \dots, \lambda_m^0)$ є точкою екстремуму функції $L(x_1, x_2, x_n, \lambda_1, \dots, \lambda_m)$.

Розглянемо метод Лагранжа для функції двох змінних

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2).$$

Таким чином, для знаходження умовного екстремуму функції $f(x_1, x_2)$ за умови $g(x_1, x_2) = 0$ необхідно знайти вирішення системи:

$$\begin{cases} L'_{x_1} = f'_{x_1}(x_1, x_2) + \lambda g'_{x_1}(x_1, x_2) = 0, \\ L'_{x_2} = f'_{x_2}(x_1, x_2) + \lambda g'_{x_2}(x_1, x_2) = 0, \\ L'_\lambda = g(x_1, x_2) = 0. \end{cases} \quad (2.104)$$

Вирішення (x_1, x_2, λ) із системи (2.104) визначає точку, в якій функція f досягає екстремуму, для цього потрібно розрахувати значення $g'_{x_1}, g'_{x_2}, L''_{x_1 x_1}, L''_{x_2 x_2}, L''_{x_1 x_2}$ і скласти визначник

$$\Delta = - \begin{vmatrix} 0 & g'_{x_1} & g'_{x_2} \\ g'_{x_1} & L''_{x_1 x_1} & L''_{x_1 x_2} \\ g'_{x_2} & L''_{x_1 x_2} & L''_{x_2 x_2} \end{vmatrix}.$$

Якщо $\Delta < 0$, то функція має в точці $(x_1^0, x_2^0, \lambda^0)$ умовний максимум, якщо $\Delta > 0$ – то умовний мінімум.

Питання і завдання для самоконтролю до розділу 2

Питання для самоконтролю:

1. У чому сутність задач лінійного програмування?
2. Які особливості задач лінійного програмування Ви можете виділити?
3. Розкрийте сутність симплексного методу.
4. Розкрийте алгоритм використання симплексного методу при вирішенні задач лінійного програмування.
5. Які методи використовуються при вирішенні задач лінійного програмування?
6. Розкрийте змістовну постановку транспортної задачі.
7. Сформулюйте математичну модель транспортної задачі.
8. Встановіть особливості вирішення закритої транспортної задачі.
9. Охарактеризуйте алгоритм визначення початкового опорного плану в транспортній задачі методом північно-західного кута.
10. Визначте напрями формування оптимального опорного плану транспортної задачі.
11. Назвіть види транспортних задач і охарактеризуйте їх.
12. Охарактеризуйте поняття «достовірність».
13. Назвіть напрями оцінки достовірності.
14. По яким напрямам відбувається аналіз лінійних моделей оптимізаційних задач?
15. Охарактеризуйте область допустимих рішень і критерій оптимальності.
16. У чому полягає інтерпретація отриманих результатів на основі лінійних оптимізаційних моделей?
17. Охарактеризуйте сутність цілочислового програмування.
18. Розкрийте напрями формулювання і вирішення задач цілочислового програмування.
19. Які методи використовуються при вирішенні задач цілочислового лінійного програмування? Охарактеризуйте їх.
20. Представте алгоритм вирішення задач цілочислового програмування.
21. У чому полягає метод Гоморі? Представте алгоритм вирішення задач цілочислового програмування цим методом.
22. У чому полягає метод віток і меж? Представте алгоритм вирішення задач цілочислового програмування цим методом.
23. Охарактеризуйте математичну модель цілочислової транспортної задачі.

24. Назвіть види і особливості вирішення задач цілочислового лінійного програмування.

25. Назвіть і охарактеризуйте основні поняття, які пов'язані з нелінійними зв'язками.

26. Визначте поняття нелінійного програмування й сутність вирішення задач нелінійного програмування.

27. Охарактеризуйте графічний метод вирішення задач нелінійного програмування при формуванні нелінійних оптимізаційних моделей.

28. Охарактеризуйте метод Лагранжа вирішення задач нелінійного програмування при формуванні нелінійних оптимізаційних моделей

Завдання для самоконтролю:

1. Підприємство випускає протягом планового періоду 2 виду продукції столи і стільці. Під час їх виробництва використовуються три види ресурсів. Дані по витратах на випуск одного виробу, запаси ресурсів, а також прибуток від реалізації одиниці продукції наведено в таблиці 2.8.

Таблиця 2.8 – Дані по витратах на випуск одного виробу, запасів ресурсів, прибутку від реалізації одиниці продукції

	Стіл	Стільці	Запас ресурсів
Ресурс 1	4	6	24
Ресурс 2	3	2	12
Ресурс 3	1	1	8
Прибуток	4	5	

Необхідно спланувати кількість виробництва столів і стільців таким чином, щоб при цих умовах виробництва прибуток був максимальним.

2. Припустимо, що у денний раціон тварин повинні входити поживні речовини двох видів в кількості, яка представлена в таблиці 2.9.

Є можливість скласти раціон із кормів двох видів, для яких задано речовини в одиницях корму і ціні однієї одиниці кожного з видів кормів.

Таблиця 2.9 – Дані про поживні речовини і вартість кормів на підприємстві

	Корм 1	Корм 2	Поживні речовини в раціоні
Поживна речовина 1	2	1	12
Поживна речовина 2	6	4	30
Ціна корму	5	2	

При задоволенні умов по необхідному змісту поживних речовин в цьому раціоні необхідно досягти його мінімальної вартості.

3. Фірма виробляє дві моделі А і В збірних книжкових полиць. Їх виробництво обмежено наявністю сировини (високоякісних дощок) і часом машинної обробки. Для кожного виробу моделі А потрібно 2 м² дощок, а для моделі В – 5 м². Фірма може одержувати від своїх постачальників до 1300 м² дощок в тиждень.

Для кожного виробу моделі А потрібно 15 хв. машинного часу, а для виробу моделі В – 30 хв. У тиждень можна використовувати 180 годин машинного часу.

Скільки виробів кожної моделі слід випускати фірмі в тиждень, якщо кожний виріб моделі А приносить 4 грн. прибутку, а кожний виріб моделі В – 2 грн. прибутку?

4. Скласти оптимальний план перевезень цегли між трьома заводами і п'ятьма об'єктами будівництва, якщо відстані (в км) між заводами і об'єктами будівництва визначаються матрицею

$$C = [C_{ij}] = \begin{pmatrix} 8 & 0 & 3 & 7 & 6 \\ 3 & 5 & 4 & 13 & 1 \\ 6 & 8 & 9 & 9 & 3 \end{pmatrix}, i = \overline{1,3}, j = \overline{1,5}.$$

Відомі потужності заводів і об'єктів будівництва.

Дані про потужність заводів і об'єктів будівництва вибираються із таблиці 2.10 і таблиці 2.11 відповідно. Варіант студент вибирає за останньою цифрою номеру залікової книжки.

Таблиця 2.10 – Потужність цеглових заводів (тис. шт. за добу)

№ заводу	Варіант									
	1	2	3	4	5	6	7	8	9	10
1	3,4	3,0	4,2	3,1	1,9	0,9	1,0	2,5	3,3	1,8
2	2,3	1,5	1,3	4,2	2,6	4,3	3,1	3,5	3,9	4,3
3	2,8	4,0	3,0	1,2	4,0	3,3	4,4	2,5	1,3	2,4

Таблиця 2.11 – Потужність об'єктів будівництва (тис. шт. за добу)

№ об'єктів будівництва	Варіант									
	1	2	3	4	5	6	7	8	9	10
1	1,5	0,9	1,1	1,0	2,7	2,1	1,2	2,5	2,8	1,3
2	1,6	2,5	1,3	3,0	1,5	2,3	1,9	2,0	1,9	0,6
3	2,1	3,0	2,2	0,6	1,0	1,4	1,8	1,7	1,1	2,4
4	1,7	0,7	3,1	1,9	3,0	1,2	1,5	0,9	0,7	3,0
5	1,6	1,4	0,8	2,0	0,3	1,5	2,1	1,4	2,0	1,2

5. Скласти оптимальний план забудови мікрорайону міста, якщо відомо, що він повинен забудовуватися житловими будинками трьох серій. Характеристики житлових будинків кожної серії подані в таблиці 2.12. З огляду на демографічний прогноз населення необхідно, щоб кількість квартир відповідала проектному завданню, що представлено в таблиці 2.13.

Дані про проектну кількість квартир вибираються із таблиці 2.13 відповідно до варіанту студента. Варіант визначається за останньою цифрою залікової книжки студента.

Таблиця 2.12 – Склад квартир і кошторисна вартість житлових будинків різних серій (для всіх варіантів однакові)

Характеристика житлових будинків	Серія		
	1	2	3
Кількість квартир - усього	200	210	150
в тому числі на двох чоловік	50	50	60
на трьох чоловік	60	70	50
на чотири чоловіки	90	90	40
Кошторисна вартість житлового будинку, тис. грн.	1200	1250	800

Таблиця 2.13 – Проектна кількість квартир у мікрорайоні на 2, 3 і 4 чоловіки

Склад сім'ї	Варіант									
	1	2	3	4	5	6	7	8	9	10
2 чол.	600	800	750	625	900	850	950	700	1000	800
3 чол.	1800	1750	1850	1750	2100	1900	2000	1850	1950	2050
4 чол.	700	650	800	600	750	550	400	850	600	450

6. Встановіть достовірність розрахунків моделі:

$$\frac{K'}{K} = 0,12 + 0,21 \cdot K_{обз} + 0,06 \cdot K_{обдз},$$

де $K_{обз}$ – коефіцієнт оборотності матеріальних запасів;

$K_{обдз}$ – коефіцієнт оборотності дебіторської заборгованості на основі встановленої погрішності.

Вихідні статистичні дані визначених показників подано в таблиці 2.14.

Таблиця 2.14 – Вихідні статистичні дані економічних показників моделі

№ спостереження	$\frac{K'}{K}$	$K_{обз}$	$K_{обдз}$
1	0,661	2,25	1,16
2	0,595	1,96	1,06
3	0,587	1,93	1,05
4	0,576	1,87	1,07
5	0,527	1,63	1,09
6	0,523	1,61	1,10
7	0,525	1,62	1,09
8	0,574	1,86	1,07
9	0,577	1,87	1,07
10	0,424	1,12	1,14
11	0,469	1,34	1,13
12	0,487	1,43	1,12
13	0,495	1,47	1,11
14	0,503	1,51	1,11
15	0,507	1,53	1,10
16	0,515	1,57	1,11
17	0,515	1,57	1,11
18	0,519	1,59	1,09
19	0,520	1,60	1,09
20	0,526	1,63	1,07
21	0,523	1,62	1,07
22	0,559	1,79	1,06
23	0,561	1,80	1,05

7. Знайти оптимальний цілочисловий план задачі.

$$Z(X) = x_1 - 3x_2 + 5x_3 + 2x_4 \rightarrow \max \text{ за умови:}$$

$$x_1 + x_2 + x_3 = 15,$$

$$2x_1 + 3x_3 + x_4 = 8,$$

$$x_j, > 0,$$

x_j – цілі числа, $j = 1, 2, 3, 4$.

8. Отримати цілочисловий оптимальний план задачі.

$$Z(X) = x_1 - 4x_2 - 2x_3 + 3x_4 \rightarrow \max \text{ за умови:}$$

$$3x_1 + x_2 + 8x_3 + x_4 = 35,$$

$$x_1 + x_3 + x_4 \leq 6,$$

$$x_j \geq 0,$$

x_j – цілі числа, $j = 1, 2, 3, 4$.

9. Контейнер обсягом 5 м^3 розміщено на контейнеровоз вантажністю 12 т. Контейнер необхідно заповнити вантажем двох найменувань. Маса одиниці вантажу m_j (в тонах), обсяг одиниці вантажу v_j (в м^3), вартості c_j (в умовних грошових одиницях) наведені в таблиці 2.15.

Таблиця 2.15 – Маса одиниці вантажу m_j (в тонах),
обсяг одиниці вантажу v_j (в м^3), вартості c_j (в умовних грошових одиницях)

Вид вантажу	m_j	v_j	c_j
1	3	1	10
2	1	2	12

Необхідно завантажити контейнер таким чином, щоб вартість вантажу, що перевозиться, була максимальною.

10. Знайти екстремуми функції:

$$L(x_1, x_2) = 2x_1 + x_2$$

при обмеженнях

$$x_1^2 + x_2^2 \leq 15, \quad x_1, x_2 \geq 0.$$

11. Підприємець вирішив виділити на розширення своєї справи 50 тис. грн. Відомо, якщо на придбання нового устаткування затрачувати x тис. грн., а на зарплату прийнятих працівників y тис. грн., то приріст обсягу продукції складе $Q = 0.001x^{0.4} \cdot y^{0.2}$.

Необхідно розподілити виділені грошові ресурси, щоб приріст обсягу продукції був максимальним.

12. Загальні витрати виробництва задані функцією:

$$T = 0,8x^2 + 0,7xy + 0,6y^2 + 800x + 500y + 1600,$$

де x і y відповідно кількість товарів A і B .

Загальна кількість виробленої продукції повинна дорівнювати 400 одиниць.

Скільки одиниць товару A і B потрібно виробити, щоб витрати на їх виготовлення були мінімальними?

РОЗДІЛ 3

УПРАВЛІННЯ РИЗИКАМИ ПІД ЧАС ОЦІНКИ НЕРУХОМОГО МАЙНА

3.1 Поняття, сутність і зміст невизначеності та ризику під час оцінки нерухомого майна

За останнє десятиріччя відбулись значні структурні зміни в економіці України. Це пов'язано із перебудовою форм власності, зміною в макро- і мікросередовищах підприємств. Невизначеність ринкових умов, зростання рівня стохастичного впливу економічних факторів призводить до зростання ризиків втрати суб'єктами підприємницької діяльності не тільки своїх позицій на ринку, але й їх занепаду.

Таким чином, в сучасних умовах більшість процесів мають невизначений характер, коли досить складно сказати, як будуть розвиватись події в майбутньому. Це ускладнює процес прийняття управлінських рішень та прогнозування показників діяльності. У таких умовах зростає рівень невизначеності й ризику. Тому важливого значення набуває визначення змісту, сутності цих понять.

Поняття «невизначеність» означає постійну зміну умов та поведінки зв'язків. Невизначеність супроводжує прийняття будь-яких управлінських рішень. Виділяються основні типи невизначеності під час прийняття управлінських рішень [41]:

- об'єктивна (природна) невизначеність;
- інформаційна невизначеність, яка характеризується недостатністю або відсутністю відповідної інформації;
- невизначеність, яка пов'язана з діяльністю суб'єктів;
- невизначеність, яка пов'язана з низькою структурованістю й ефективністю організаційної структури;
- невизначеність, що обумовлена нечіткістю, розпливчатістю процесів, явищ, інформації.

Ухвалення рішень в умовах невизначеності є вибором тієї або іншої альтернативи з їх різноманіття, а сам процес ухвалення рішень нерозривно пов'язаний з перетворенням невизначеності у визначеність.

Невизначеність обумовлюється неповнотою, несвоєчасністю, низьким рівнем специфікації, суперечливістю інформації [42].

В умовах зростання рівня невизначеності відповідно збільшується ризик діяльності підприємств.

Слід вказати, що залежність між рівнем невизначеності економічної ситуації і рівнем ризику в оцінці нерухомого майна досить висока. Чим більше невизначеність, тим більше ризик і навпаки. Тому необхідно обґрунтувати поняття ризику, його сутність, оцінити його рівень і встановити причини його виникнення.

Термін «ризик» походить від грецьких слів *ridsikon*, *ridsa* – круча, скала.

Ризик – це суб'єктивна характеристика невизначеності сценарію реалізації оціночного проекту [43].

Ризик – це об'єктивно-індивідуальна категорія, яка характеризує загрозу, небезпеку виникнення негативних явищ в економічному та організаційному середовищі.

Ризик – це можливість відхилення від мети задля досягнення рішення; вірогідність помилки або успіху того або іншого вибору в ситуації з деякими альтернативами; це ситуативна характеристика діяльності, що полягає в невизначеності її результату в можливих несприятливих наслідках у разі «неуспіху»; рівень невизначеності в прогнозі результату [44].

Ризик – вибір управляючих параметрів (управляючих дій), що не гарантує виконання поставленої мети у зв'язку з невизначеністю [45] умов господарювання.

Ризик – це діяльність, пов'язана з подоланням невизначеності в ситуації неминучого вибору, в процесі якої є можливість кількісно і якісно оцінити вірогідність досягнення передбачуваного результату невдачі і відхилення від мети [46].

Як видно з теоретичних визначень «ризик», їх об'єднує невизначеність здійснення події в позитивну або негативну сторону. В цьому аспекті виникає необхідність оцінки ризику для мінімізації негативних явищ і забезпечення позитивного результату.

Сутність ризику полягає в співвідношенні мети і результату діяльності, між якими існують відхилення. Ступінь цього відхилення і визначає рівень ризику діяльності. Якщо відхилення мети і результату зростає, то і збільшується ризик, якщо навпаки – то останній зменшується [47].

В умовах реформування економіки управління ризиком є складним процесом, який включає багато чинників. Управління ризиком складається з чотирьох блоків [48].

Перший блок – це порівняльна характеристика ризику. Суть його полягає в розрахунку кількісних показників, показники ризику порівнюються із

стандартними величинами ризику, інформацією нормативних документів або порівнянними показниками ризиків.

В другому блоці визначається наскільки значний наявний ризик.

В третьому блоці відбувається моніторинг ризику та здійснення управлінських дій щодо його зниження.

В четвертому блоці здійснюється контроль за ризиком і розповсюдження інформації про його рівень.

Управління ризиком є важливим і постійним елементом управління, який впливає із загальних концепцій ризику [49]:

- перша – «оборонна», пасивна, яка реалізується за допомогою страхування і відокремлення від ризику;

- друга – «наступальна», активна, реалізована в ході зростання, оновлення, накопичення тієї інформації, яка необхідна для скорочення області невизначеності і ризику.

Виділяють такі етапи управління ризиком [48]:

- 1 етап – формування цілей у відповідній сфері діяльності;

- 2 етап – визначення критеріїв оцінки різних заходів по управлінню ризиком;

- 3 етап – вибір і аналіз заходів по досягненню визначених цілей;

- 4 етап – вибір найбільш адекватних заходів і контроль результатів їх виконання.

На першому етапі визначається масштаб рівня ризику, суб'єкти і об'єкти ризику, спрямованість заходів щодо зниження ризиків, формуються задачі.

На другому етапі вибирають оптимальні заходи щодо скорочення або запобігання ризиків, які б були спрямовані на їх досягнення. До початку вибору заходів щодо управління ризиком потрібно вирішити, за допомогою яких критеріїв оцінювати вибрані заходи. Необхідною умовою вибору переважної більшості заходів щодо управління ризиком є їх технологічна і економічна здійснимість.

На третьому етапі аналізуються заходи щодо управління ризиком.

На четвертому етапі проводиться постійний контроль виконання різних етапів робіт.

При проведенні контролю результатів робіт треба враховувати [50]:

- відповідність вибраних показників контролю меті і задачам здійснюваних заходів;

- проведення контролю результатів робіт через певний проміжок часу;

- проведення контролю всіх заходів (експозиції, ризику, економічних показників і т.п.);
- використання показників відповідного масштабу.

Аналіз заходів щодо управління ризиком спрямований на досягнення основної мети їх діяльності – забезпечення розвитку, і складається з декількох етапів. На першому етапі в процесі управління ризиком використовують систему показників або «набір інструментів», які спрямовані на зниження ризику або формують підходи щодо запобігання появи ризику. Тобто важливо не констатувати появу ризику, а його діагностувати. Це дозволить встановлювати причинно-наслідкові зв'язки виникнення та розвитку ризику в оцінці нерухомого майна.

На другому етапі аналізують запропоновані заходи, щоб оцінити переваги різних стратегій. Найефективніші заходи можна потім більш ретельно проаналізувати, щоб вибрати найприйнятніші з них. На цьому етапі особливо важливе залучення широкого круга зацікавлених сторін, щоб вибрані стратегії не викликали негативного відношення.

На наступному етапі розробляється система дій щодо впровадження стратегії управління ризиком [51]. Цей етап передбачає розробку конкретних заходів і підходів управління ризиком в конкретних умовах.

На останньому етапі впроваджується стратегія і здійснюється її моніторинг.

Таким чином, аналіз заходів управління ризиком – це система дій, яка спрямована на зниження рівня ризику або його усунення для досягнення основної мети в оцінці нерухомого майна.

3.2 Система показників кількісного оцінювання ступеня ризику в оцінці нерухомого майна

Оцінка ризику – це систематичний процес виявлення факторів і видів ризику та їх кількісну оцінку, тобто методологія аналізу ризиків поєднує кількісний і якісний підходи [52].

Оцінка ризику здійснюється в два етапи: якісний і кількісний [53]. У рамках першого етапу встановлюються джерела й причини ризику, етапи або роботи, при виконанні яких виникає ризик, тобто визначаються потенційні зони ризику, виявляються ризики, які пов'язані з діяльністю підприємства, прогнозуються практичні вигоди і можливі негативні наслідки ризиків. Результати якісного аналізу є важливими для проведення кількісного аналізу.

На другому етапі розраховуються кількісні значення окремих ризиків і ризиків підприємства в цілому. Також виявляється можливий збиток і дається вартісна оцінка від прояву ризику і, нарешті, завершальною стадією кількісної оцінки є вироблення системи антикризових заходів і розрахунок їх вартісного еквівалента.

Розглянемо більш детально особливості кількісного оцінювання рівня ризику [54].

Кількісне оцінювання рівня ризику – це важливий етап процесу управління, який має включати оцінювання реального (фактичного) ризику, а також встановлення меж допустимого ризику для окремих господарських операцій, організаційних підрозділів і фінансових установ. Водночас потрібно оцінити й ризики освоєння нових ринків, продуктів і напрямів діяльності. Ризик економічних рішень оцінюється сподіваними втратами, що є наслідками цього рішення. Ступінь ризику вимірюється втратами (збитками), яких можна очікувати в разі реалізації цього ризику, а також ймовірністю, з якою ці втрати можуть статися. Коли ймовірність втрат висока, а розмір їх малий або навпаки – збитки малоймовірні, хоча й оцінюються як суттєві, то ризик вважається невисоким (малим). Методи оцінки ризику під час оцінки нерухомого майна, що формалізують процес вимірювання та розрахунків, мають визначати три основні компоненти ризику [55]:

- розмір (величина) – сума можливих втрат;
- ймовірність настання негативної події;
- тривалість періоду впливу ризику.

Слід зазначити, що кількісна оцінка ризику визначається величиною можливих ймовірнісних втрат, тому необхідно враховувати випадковий характер таких втрат. Ймовірність настання події може бути визначена об'єктивним і суб'єктивним методами.

Перша група методів використовується для визначення ймовірності настання події на основі розрахунку частоти, з якою відбувається ця подія. Це методи теорії ймовірностей, економічної статистики, теорії ігор та інші математичні методи.

Суб'єктивні методи базуються на використанні суб'єктивних критеріїв, які базуються на різних припущеннях. До таких припущень відносяться судження оцінюючого, його особистий досвід, оцінка експерта по рейтингу та інше [55]. Суб'єктивні методи застосовують тоді, коли ризики не піддаються кількісному вимірюванню – квантифікації.

Для оцінки величини фінансових ризиків в основному використовують три групи показників [56]:

- статистичні величини (стандартне відхилення, варіація, дисперсія, коефіцієнт бета);
- непрямі показники ризикової діяльності, обчислені, зазвичай, у формі фінансових коефіцієнтів за даними публічної звітності;
- аналітичні показники (індикатори), призначені для оцінки конкретного виду ризику (валютного, відсоткового, кредитного, інвестиційного, незбалансованої ліквідності) в процесі внутрішнього аналізу діяльності підприємницьких структур.

Для мінімізації або нівелювання впливу ризиків можуть бути використані кумулятивні методи. Ці методи базуються на експертних оцінках, пов'язаних з інвестуванням. При використанні кумулятивного методу на першому етапі розраховується безризикова ставка доходу. Наступним етапом є оцінка надбавки, яка додається до безризикової ставки доходу. Таким чином, кількісно ставка дисконтування розраховується:

$$C_d = \text{БСД} + \Sigma \text{IP}, \quad (3.1)$$

де C_d – ставка дисконту;

БСД – безризикова ставка доходу;

ΣIP – сума інвестиційних ризиків.

В цілому кумулятивний метод визначається як один із методів оцінки коефіцієнта (ставки) капіталізації, коли коефіцієнт складається з декількох складових (безризикова ставка; премія за ризик; премія за низьку ліквідність; премія за управління інвестиціями; фактор фонду відшкодування).

Під час використання цього методу запропоновано таке співвідношення:

$$Y = 26,41 \ln X - 64,16, \quad (3.2)$$

де Y – базова ставка дисконту;

X – діюча ставка Центрального банку.

Запропонована формула дозволяє конкретизувати кількісну характеристику оцінки інвестиційних ризиків. Проте, отримане в результаті кореляційно-регресійного аналізу співвідношення, базується на конкретних показниках минулих періодів. Тому воно носить обмежений характер і віддзеркалює лише ті умови, в яких формувались ці показники.

Слід зазначити, що підходи щодо впровадження кумулятивних методів тісно пов'язані також з оцінкою премії за ризик для конкретного підприємства.

Премія за ризик диференціюється в залежності від об'єкту інвестування, або від умов інвестування реалізації проекту. Першими факторами виступають: використання нової технології, обладнання, нової техніки, випуск нової продукції та інші. До других факторів відносять: попит і пропозицію, вплив зовнішньоекономічних ризиків, внутрішньо економічні ризики, політична нестабільність та інші.

Досить проблемним питанням є визначення сумарного ризику. В цьому аспекті показник ризику може бути кількісно оцінений таким чином:

$$\Sigma P = (1+a_1) * \dots * (1+a_k) - 1, \quad (3.3)$$

де ΣP – сумарний ризик;

$a_1 \dots a_k$ – види ризиків.

В умовах інфляційних процесів ставка дисконту визначається співвідношенням:

$$R_n = R_p + i + R_{pxi}, \quad (3.4)$$

де R_n – номінальна ставка дисконту;

R_p – реальна ставка дисконту;

i – темп інфляції.

Запропоновані підходи щодо визначення категорії «ризик» і напрями його оцінки дозволять підприємствам своєчасно реагувати на негативні явища, які пов'язані з макросередовищем і внутрішньогосподарською діяльністю. Слід вказати, що прийняття ефективних управлінських рішень базується на кількісному підґрунті. В цьому аспекті представлені методи кількісної оцінки ризику, що дозволяють оцінити ці ризики підприємства і встановити рівень премії за них.

Система кількісних оцінок ризику в абсолютному виразі складається з таких [57]:

- у випадку, коли рішення є альтернативним, тобто можливі лише два наслідки його реалізації, показники ризику розраховуються за такою залежністю:

$$R = X_n P_n, \quad (3.5)$$

де X_n – величина збитків у разі настання негативного наслідку рішення;

P_n – ймовірність настання негативного наслідку;

- у випадку, якщо рішення мають декілька (множину) наслідків реалізації, використовують такі показники:

1) математичне сподівання. Математичне сподівання дискретної величини є сумою добутків можливих варіантів цієї величини на їх імовірність:

$$M(x) = \sum_{i=1}^n X_i \cdot P_i, \quad (3.6)$$

причому основною умовою використання цієї формули є:

$$\sum_{i=1}^n P_i = 1. \quad (3.7)$$

Математичне сподівання для неперервної величини:

$$M(x) = \int_{-\infty}^{+\infty} f(x) dx; \quad (3.8)$$

2) показник дисперсії характеризує ступінь мінливості реальних даних деякої випадкової величини навколо математичного сподівання. Визначається як математичне сподівання квадратів відхилень індивідуальних значень випадкової величини від її математичного сподівання:

$$\sigma^2 = M(x - M(x))^2. \quad (3.9)$$

Для дисперсійної величини формула дисперсії має вигляд:

$$\sigma^2 = \sum_{i=1}^n (x_i - M(x))^2 \cdot P_i. \quad (3.10)$$

Для безперервної величини:

$$\sigma^2 = \int_{-\infty}^{+\infty} M(x - M(x))^2 f(x) dx; \quad (3.11)$$

3) середньо квадратичне відхилення:

$$\sigma = \sqrt{M(x - M(x))^2}, \quad (3.12)$$

$$\sigma = \sqrt{\sum_{i=1}^n (x_i - M(x))^2 \cdot P_i}. \quad (3.13)$$

Іноді для оцінки величини ризику в абсолютному виразі використовують ймовірність настання небажаних наслідків, тобто величини P .

Для оцінки ризику при обґрунтуванні управлінських рішень, в більшості випадків, не достатньо використання абсолютних показників. В системі оцінки ризику використовують відносні показники.

У відносному виразі ризик визначається коефіцієнтом ризику, який визначається як відношення величини максимальних втрат від даного виду діяльності до деякої бази порівнянь (за таку базу може прийматись обсяг власних ресурсів підприємства, загальні величини втрати по даному виду діяльності або сподіваний дохід від даного виду діяльності):

$$K_p = X/K, \quad (3.14)$$

де X – величина максимально можливих втрат;

K – база порівнянь.

Цей показник, як правило завершує проведення дисперсійного аналізу ризику і використовується при наявності масиву статистичної інформації. Причому, чим більший цей показник, тим більшим є ризик, пов'язаний з даним проектом.

В системі оцінки ризику необхідно визначити границі або інтервали, де можна допускати відповідний рівень ризику, де він є критичним. Тобто визначають *зони ризику* [58]. При цьому використовують графічний аналіз і аналітичні підходи щодо визначення цих зон.

Графічний аналіз полягає в побудові кривої щільності розподілу ймовірностей настання ризику [59]. При цьому визначаються зони допустимого й критичного ризику. Наприклад, для інвестиційного проекту будують криву, що висвітлює залежність між ймовірними збитками по проекту і сподіваними віддачами цього проекту. У цьому процесі визначають точки, які відповідають зонам критичного або допустимого ризику.

Точка допустимого ризику характеризує найбільш ймовірні збитки по проекту і сподівану або середню віддачу цього проекту, в якій збитки будуть мати величину, що дорівнює загальній величині прибутку від проекту [60].

Ця точка є верхньою межею зони допустимого ризику.

Ймовірність допустимого ризику визначається залежністю:

$$F(x) = \int_0^{X_{доо}} f(x) dx. \quad (3.15)$$

Під зоною допустимого ризику розуміють область, у межах якої відповідний вид діяльності зберігає свою економічну доцільність, тобто випадкові збитки не перевищують очікуваного ризику від проекту [61].

Точка допустимого критичного ризику характеризує ступінь гранично допустимого критичного ризику, тобто ризику втрат, що досягає величини розрахункової виручки від проекту [62].

Ймовірність критичних ризиків визначається залежністю:

$$F(x) = \int_{X_{доо}}^{X_{кр}} f(x) dx. \quad (3.16)$$

Під зоною критичного ризику розуміють область випадкових збитків, розміри яких перевищують величину очікуваного збитку і сягають величини розрахованої виручки [62].

Точка катастрофічного ризику характеризує ступінь гранично-катастрофічного ризику, тобто ризику втрат, що сягає розміру всього майна підприємства [62].

Ймовірність катастрофічного ризику визначається шляхом інтегрування:

$$F(x) = \int_{X_{кр}}^{X_{кат}} f(x) dx. \quad (3.17)$$

Зона катастрофічного ризику – це область можливих втрат, які перевищують величину розрахованої виручки і можуть сягати вартості майна підприємця [63].

Катастрофічний ризик може призвести підприємство до банкрутства, крім того, до катастрофічних відносяться всі ризики пов'язані із загрозою для життя людей, оточуючого середовища, тощо.

Частіше за все під час прийняття рішень підприємця цікавить не стільки ймовірність певного рівня втрат, скільки ймовірність, що його втрати не перевищать певної позначки

$$W(x) = 1 - F(x), \quad (3.18)$$

де $W(x)$ – це функція розподілу ймовірностей перевищення певного рівня випадкових збитків.

Відповідно до визначених зон ризику розраховують такі показники ризику [64]:

- *показник допустимого ризику* – це ймовірність того, що втрати виявляться більшими за гранично допустимий рівень (таким рівнем є прибуток від проекту);
- *показник критичного ризику* – це ймовірність того, що втрати виявляться більшими за допустимий критичний рівень (розрахункова виручка);
- *показник катастрофічного ризику* – це ймовірність того, що втрати по проекту виявляться більшими за граничний катастрофічний рівень (вартість майна підприємця).

Таким чином, представлені показники дають змогу комплексно оцінити стадії ризику проекту й прийняти управлінські рішення, які будуть мати мінімальний ризик.

Відповідно для цих показників визначають і граничні значення допустимого, критичного або катастрофічного ризиків. Граничні значення представлено в таблиці 3.1.

Таблиця 3.1 – Граничні значення

Показник	Значення показника
Граничне значення допустимого ризику	0,1
Граничне значення критичного ризику	0,01
Граничне значення катастрофічного ризику	0,001

На основі граничних значень показників ризику приймаються управлінські рішення. Так, граничне значення допустимого ризику не повинно перевищувати значення 0,1, граничне значення критичного ризику – 0,01, а граничне значення катастрофічного ризику – 0,001.

Підприємства на кожному етапі господарської діяльності здійснює відповідні інвестування грошових коштів в економічний процес. Тому необхідно постійно моніторити цей процес, виявляти негативні явища й встановлювати рівень ризику ліквідності. Потреба в оцінці ризику ліквідності виникає і під час змін стратегії й тактики діяльності підприємства.

Ризик ліквідності – це форма ризику, яка показує ймовірність погашення зобов'язань підприємством на кожному етапі інвестування грошових коштів у виробничий процес [64]. Цей ризик пов'язаний з низьким рівнем віддачі об'єктів інвестування, неефективним створенням відповідних зобов'язань, відсутністю необхідного розміру грошових коштів і т.п.

Для оцінки ризику ліквідності використовують два критерії [62]:

- період переходу інвестицій у грошові кошти, які в залежності від часу їх трансформації бувають:

- 1) терміново ліквідні з незначним ризиком (час трансформації до 7 днів);
- 2) високоліквідні інвестиції з низьким ризиком (час трансформації від 7 до 30 днів);
- 3) середньо ліквідні із середнім ризиком (час трансформації від 1 до 3 місяців);
- 4) мало ліквідні об'єкти з високим ризиком (час трансформації більше 3 місяців).

Чим більшим є показник ризику ліквідності, тим меншим є ризик ліквідності;

- оцінка ліквідності інвестицій за рівнем фінансових витрат здійснюється на основі розрахунку процентного співвідношення величини можливих втрат до обсягів інвестицій, які прагнуть реалізувати.

За цим критерієм всі об'єкти інвестування оцінюють як: з дуже високим ризиком (витрати перевищують 20%); з високим ризиком (11-20%); із середнім ризиком (6-10%); з низьким ризиком (до 5%).

Слід вказати, що існує залежність між показниками ризику ліквідності за періодом і рівнем фінансових втрат. Якщо зростає показник ризику ліквідності, то спостерігається скорочення оцінки ризику ліквідності інвестицій за рівнем фінансових втрат і, навпаки. Це пов'язано з тим, що при здійсненні інвестицій у виробничий процес його суб'єкти хочуть швидше реалізувати і отримати віддачу від проекту за короткий термін при досить великому рівні фінансових втрат.

Таким чином, при здійсненні виробничого процесу представлені напрями оцінки ризику дають змогу виявляти його, моніторити на кожному етапі процесу, здійснювати контроль за його рівнем і не тільки констатувати факт виникнення ризику, а й діагностувати й приймати управлінські рішення.

Питання і завдання для самоконтролю до розділу 3

Питання для самоконтролю:

1. Назвіть типи невизначеності в задачах ухвалення управлінських рішень.
2. Визначте категорію «ризик».
3. Охарактеризуйте аспекти управління ризиком.
4. Назвіть і охарактеризуйте етапи управління ризиком.
5. Назвіть основні напрями аналізу при здійсненні управління ризиком.
6. У чому полягає кількісна оцінка ризику?
7. Які показники використовуються для кількісної оцінки ризику?
8. Охарактеризуйте систему кількісних оцінок ризику в абсолютному виразі.
9. Охарактеризуйте систему показників визначення ризику у відносному виразі.
10. Визначте напрями оцінки допустимого і критичного ризику.
11. Охарактеризуйте напрями оцінки ризику ліквідності.

Завдання для самоконтролю:

Найдіть правильну відповідь у таким тестових завданнях:

1. Основні типи невизначеності в задачах при прийнятті управлінських рішень:
 - А) інформаційна невизначеність, яка характеризується недостатністю або відсутністю відповідної інформації;
 - Б) організаційна невизначеність;
 - В) математична невизначеність;
 - Г) об'єктивна (природна) невизначеність.
2. Ризик – це:
 - А) негативні явища, які виникають в обумовленій системі;
 - Б) можливість відхилення від мети, заради досягнення якої ухвалювалося рішення;

В) система дій, яка спрямована на розвиток підприємства;

Г) вибір управляючих параметрів (управляючих дій), що не гарантує виконання поставленої мети у зв'язку з невизначеністю (характером вірогідності) умов господарювання.

3. Якому етапу відповідають такі дії:

1 етап _____

2 етап _____

3 етап _____

4 етап _____

А) визначення критеріїв оцінки різних заходів по управлінню ризиком;

Б) вибір найбільш адекватних заходів і контроль результатів їх виконання;

В) формування цілей у відповідній сфері діяльності підприємства;

Г) вибір і аналіз заходів по досягненню визначених цілей.

4. Розташуйте етапи аналізу заходів управління ризиком у відповідності до їх здійснення:

1 етап _____

2 етап _____

3 етап _____

4 етап _____

5 етап _____

А) використовують систему показників або «набір інструментів», які спрямовані на зниження ризику або формують підходи щодо запобігання появи ризику;

Б) залучення експертів щодо вибору стратегії управління ризиком;

В) аналіз запропонованих заходів, щоб оцінити переваги різних стратегій;

Г) впровадження стратегії і здійснення її моніторингу;

Д) розробка системи дій щодо впровадження стратегії управління ризиком.

1. Емігрант з України включається в гру на фондовій біржі, після того як отримав роботу і має стабільний дохід. Заощадивши власні 10000 доларів, він взяв у борг ще 40000 доларів під 10 % річних і вклав всі 50000 доларів в акції однієї з компаній, розраховуючи на річне зростання курсу 20 %. Але фактичний курс почав падати з ряду причин і коли він знизився на 40 % емігрант вирішив

позбутися ненадійних акцій, у результаті чого збитки призвели його до банкрутства. Його знайомий американець також вклав власні 50000 доларів в акції тієї ж фірми, а потім продав їх, проте американцю вдалося уникнути банкрутства. Чому збанкрутував емігрант?

2. Необхідно інвестувати тимчасово вільні грошові кошти строком на 2 роки з тим, щоб в кінці отримати суму рівну 1260000 грн. На ринку доступні 2 види фінансових інструментів – дисконтні облігації терміном звернення 1 рік і 3 року (номінальна вартість 126 грн.). Поточна ціна річних облігацій складає 100.8 грн., трирічних – 64.5 грн. Прибутковість як одного, так і іншого виду облігацій складає 25 %.

Визначити необхідну суму інвестицій при незмінності ставок прибутковості протягом всього терміну інвестування.

РОЗДІЛ 4

ЕКОНОМЕТРИЧНЕ МОДЕЛЮВАННЯ В ОЦІНЦІ НЕРУХОМОГО МАЙНА

4.1 Принципи побудови економетричних моделей. Парна лінійна регресія

У математичному моделюванні важливе місце займають економетричні моделі, які дозволяють встановити причинно-наслідковий зв'язок між економічними факторами [65]. На основі економетричних моделей розробляються організаційно-економічні механізми діяльності підприємства, формуються управлінські рішення, які кількісно й якісно відображають економічні процеси, що відбуваються в сфері діяльності суб'єктів підприємницької діяльності [66]. Слід також зазначити, що більшість економічних процесів мають стохастичний невизначений характер. Стохастична залежність може бути суттєвою, тобто обумовлена внутрішніми властивими даному явищу причинами, і несуттєвими, які викликані дією зовнішніх (випадкових) причин (середовищем); безпосередньою, стійкою і нестійкою, сильною і слабкою, простою (між двома змінними) і складною (між залежною змінною Y і декількома чинниками-аргументами x_1, x_2, \dots, x_n). Для оцінки саме таких процесів і використовують методи економетричного моделювання.

Економетричні моделі бувають [67]:

1. Парними – відображають причинно-наслідковий зв'язок між незалежним фактором (x) і залежною змінною (y). Наприклад, модель залежності між середньо списковою чисельністю працюючих ($Ч$) і рентабельністю витрат (P_e):

$$P_e = 0,07 + 0,18 Ч, \quad (4.1)$$

модель залежності між рентабельністю реалізації продукції (P_n) і рівнем витрат (P_{en}):

$$P_n = 0,018 - 0,21 P_{en}. \quad (4.2)$$

2. Багатофакторними – відображають причинно-наслідковий зв'язок між декількома незалежними факторами (x) і залежною змінною (y).

$$P_e = 0,23 + 0,25 Ч - 0,16 P_{en}. \quad (4.3)$$

Розглянуті моделі відображають залежність між економічними факторами, що характеризуються математичною формою:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n.$$

Слід зазначити, що розглянуті моделі висвітлюють лінійну форму залежності між показниками. Проте більшість економічних процесів мають нелінійний характер. Тому для спрощення проведення економетричного дослідження необхідно використовувати методи лінеаризації для переходу від нелінійної форми до лінійної.

Таким чином, *економетрична модель* – це кількісне відображення причинно-наслідкових зв'язків між економічними факторами, яке створює підґрунтя для побудови організаційно-економічних механізмів управління економічними процесами підприємства.

В економетричних моделях незалежні змінні x_1, x_2, \dots, x_n називають *пояснюючими змінними* (або *факторами*, інколи регресорами) [68]. Залежні змінні y називають *пояснюваними змінними* (або регресандами). Усі змінні економетричних моделей, як і будь-якої математичної моделі, поділяють на *екзогенні* і *ендогенні* [68].

Екзогенними (зовнішніми) називаються змінні, значення яких є наперед визначеними перед використанням моделі, а *ендогенними* (внутрішніми) – такі, значення яких визначаються тільки із самої моделі [68].

В економетричному моделюванні необхідно визначити принципи побудови цих моделей. Принцип від латинської *principium* – основа, початок.

Існують декілька визначень категорії «принцип» [69]:

1. Основні положення передумови.
2. Теоретичне знання, що не є ні доказовим, ні вимагаючим доказу.
3. Етична норма, яка, згідно Канту, є або суб'єктивною або об'єктивною.

Виходячи з вище сказаного, основними принципами побудови економетричних моделей є [70]:

1. Інформація, що використовується в економетричному моделюванні, повинна бути повною, достовірною, що адекватно відображає економічні процеси, які відбуваються на підприємстві.
2. Фактори, що включаються в економетричну модель, повинні бути кількісно оцінені, економічно інтерпретовані, кількість спостережень на один показник складає не менше 8.

3. Математичний апарат, який використовується в економетричному моделюванні, повинен вирішити проблему побудови моделі і за його допомогою можна оцінити адекватність цієї моделі.

4. Під час розробки економетричної моделі необхідно враховувати випадковий член. Випадковий член існує за декількома причинами:

- не включення пояснювальних змінних. Співвідношення між y і x завжди є досить великим спрощенням. У дійсності існують інші фактори, які здійснюють вплив на y , і які не включені в економетричну модель;

- агрегування змінних. У багатьох випадках залежність, що розглядається, це спроба об'єднати разом деяке число економічних співвідношень;

- помилковий опис структури моделі. Структура моделі може бути описана помилково;

- функціональна специфікація може бути помилково визначена;

- помилково проведені розрахунки в моделюванні.

5. Параметри моделі повинні економічно адекватно відображати причинно-наслідковий зв'язок між факторами. Наприклад, в моделі (4.1) зростання середньо спискової чисельності працівників на 1 робітника призведе до збільшення рентабельності витрат на 18 коп./грн.

6. Економетрична модель повинна бути перевірена на адекватність шляхом використання відповідних критеріїв.

Для оцінки зв'язку між факторами економетричної моделі використовують критерії: коефіцієнт кореляції і коефіцієнт детермінації.

Коефіцієнт кореляції показує ступінь впливу незалежного фактору (x) на залежну змінну (y). Цей критерій використовується в парних економетричних моделях – коефіцієнт парної кореляції, і в багатофакторних економетричних моделях – коефіцієнт множинної кореляції.

Коефіцієнт кореляції показує, на яку частину середнього квадратичного відхилення змінюється функція y , якщо аргумент x збільшується (зменшується) на своє середньоквадратичне відхилення σ_x [71]. Знак коефіцієнта парної кореляції співпадає із знаком коефіцієнта регресії, а його чисельне значення коливається в межах:

$$-1 \leq r_{y/x} \leq 1. \quad (4.4)$$

Коефіцієнт парної кореляції може бути визначений таким чином [71]:

$$r_{y/x} = \frac{\overline{XY} - \bar{X} \times \bar{Y}}{\sigma_y \times \sigma_x}, \quad (4.5)$$

де $r_{y/x}$ – коефіцієнт парної кореляції;

\bar{X} – середнє значення незалежної змінної X ;

\bar{Y} – середнє значення залежної змінної Y ;

σ_x – середнє квадратичне відхилення показника X ;

σ_y – середнє квадратичне відхилення показника Y .

Слід зазначити, що середнє квадратичне відхилення визначається за формулою:

$$\sigma_y = \sqrt{D_y}, \quad (4.6)$$

де D_y – дисперсія (середній квадрат відхилення).

Дисперсія визначається:

$$D_y = \overline{Y^2} - \bar{Y}^2, \quad (4.7)$$

де $\overline{Y^2}$ – середній квадрат показника Y ;

\bar{Y}^2 – квадрат середнього для показника Y .

Аналогічні формули використовуються і для показника x .

Розглянемо приклад. Розрахуйте коефіцієнти кореляції і детермінації на основі представлених у таблиці 4.1 спостережень.

Таблиця 4.1 – Таблиця вихідних даних для проведення розрахунків

Спостереження	x	y
1	1	3
2	2	5
3	3	6
Сума	6	14
Середнє	2	4,667

Розв'язання

Визначимо дисперсію для факторів y і x . Для цього складемо таблицю 4.2.

Таблиця 4.2 – Розрахунок середніх значень показників y і x

Спостереження	x	y	x^2	y^2
1	1	3	1	9
2	2	5	4	25
3	3	6	9	36
Сума	6	14	14	70
Середнє	2	4,667	4,667	23,33
Квадрат середнього	4	21,78		

Використовуючи формулу (4.7) визначимо дисперсію для y і x :

$$D_y = 23,33 - 21,78 = 1,55,$$

$$D_x = 4,667 - 4 = 0,667.$$

Визначимо середнє квадратичне відхилення показників y і x , використовуючи формулу (4.6):

$$\sigma_y = \sqrt{1,55} = 1,24,$$

$$\sigma_x = \sqrt{0,667} = 0,82.$$

Використовуючи формулу (4.5), визначимо коефіцієнт парної кореляції для y і x :

$$r_{y/x} = \frac{10,333 - (2 \times 4,667)}{1,24 \times 0,82} = \frac{0,999}{1,017} = 0,98.$$

Значення коефіцієнту парної кореляції, що характеризують силу впливу показника x на y подано в таблиці 4.3.

Таблиця 4.3 – Значення коефіцієнта парної кореляції

Значення коефіцієнта кореляції	Сила впливу показника x на y
0,85 - 1	сильний
0,55 - 0,84	помірний
0,25 - 0,54	слабкий
0 - 0,24	дуже слабкий

Знак значення коефіцієнта парної кореляції вказує на напрямок зв'язку. Якщо знак «+», то це вказує на прямо пропорційний зв'язок між факторами, якщо навпаки – то обернений.

Коефіцієнт множинної кореляції використовується в багатфакторному економетричному аналізі. Його значення знаходиться між 0 і 1. Сила впливу показників x на результуючий фактор y характеризується значеннями поданими в таблиці 4.3.

Коефіцієнт детермінації визначається як квадрат коефіцієнту кореляції:

$$D_{y/x} = r_{y/x}^2$$

або

$$D_{y/x_i} = r_{y/x_i}^2.$$

Основними напрямками оцінки адекватності економетричної моделі є:

1. Перевірка за допомогою F-тесту (F-критерій Фішера).
2. Використання t-розподілу Ст'юдента для оцінки надійності коефіцієнта кореляції.
3. Перевірка моделі на гомо-гетескедастичність.
4. Перевірка факторів економетричної моделі на мультиколінеарність.

F-тест використовується для оцінки того чи важливе пояснення, яке дає рівняння в цілому. Цей тест заснований на порівнянні залишкової теоретичної дисперсії $\tilde{\sigma}_{y/x}^2$ і загальної дисперсії σ_y^2 . Розглядається відношення

$$\frac{\sigma_{y/x}^2}{\sigma_y^2} = T_{розр} \text{ і порівнюється з табличним (для } \bar{f} \% \text{ Фішера знайдено розподіл}$$

і складена спеціальна таблиця) при заданому рівні значущості і різних ступенях свободи.

Загальна дисперсія σ_y^2 досліджених даних від їх середнього значення встановлюється з урахуванням числа ступеней свободи $f = n - k$:

$$\sigma_{y/x}^2 = \frac{\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y})^2 m_i}{n - k}, \quad (4.8)$$

де k – число інтервалів у вибіркових даних.

Залишкова теоретична дисперсія $\tilde{\sigma}_{y/x}^2$ встановлюється як різниця розрахункових \tilde{y}_i і середніх інтервальних значень \bar{y}_i з урахуванням числа ступеней свободи:

$$d_1 = K - P \text{ і } d_2 = n - K,$$

де P – число параметрів управління.

Якщо $\bar{f}_{\text{розн}} \leq \bar{f}_{\text{табл}}$, то при заданому рівні значущості складене рівняння регресії затверджується. Вірогідність помилки тим менша, чим більше рівень значущості α %.

У разі, коли чисельник $\tilde{\sigma}_{y/x}^2$ менше за знаменника σ_y^2 , то міняємо їх місцями разом з відповідними ступенями свободи:

$$d_1 = K - P \text{ і } d_2 = n - K.$$

Розглянемо приклад. Загальна дисперсія $\sigma_y^2 = 41,5$ при $n = 154$ і $K = 12$.

Залишкова дисперсія $\tilde{\sigma}_{y/x}^2 = 34,44$ при $K = 12$ і $P = 3$ ($P = 3$ в квадратному рівнянні регресії).

Розв'язання

$$T_{\text{розн}} = \frac{\tilde{\sigma}_{y/x}^2}{\sigma_y^2}, \text{ оскільки } \sigma_y^2 > \tilde{\sigma}_{y/x}^2, \text{ переходимо до відношення із}$$

ступенями свободи:

$$d_1 = 154 - 12 = 142; d_2 = 12 - 3 = 9,$$

$$\bar{f}_{\text{розн}} = \frac{41,5}{34,44} = 1,21,$$

за таблицю $\bar{f}_{5\%}(142,9) = 2,75$ та $\bar{f}_{20\%}(142,9) = 1,7$.

Отже, знайдене квадратне рівняння регресії з високою надійністю узгоджується з вихідними даними.

Слід зазначити, що в регресійному аналізі побудова F -статистики здійснюється шляхом розподілення дисперсії залежної змінної на «пояснювальні» і «непояснювальні» складові [72]:

$$F = (ESS / k) / RSS / (n - k - 1), \quad (4.9)$$

де ESS – пояснювальна сума квадратів відхилень;

RSS – залишкова (непояснювальна) сума квадратів;

k – кількість ступенів свободи;

n – кількість значень факторів моделі.

При здійсненні F-тесту для рівняння перевіряється чи перевищує r^2 те значення, яке може бути отримано випадково. Для розрахунку F-статистики для рівняння в цілому формулу (4.9) можна трансформувати шляхом ділення чисельника і знаменника рівняння на TSS (загальну суму квадратів), відмічаючи, що ESS/TSS дорівнює r^2 , а RSS/TSS дорівнює $(1-r^2)$.

У результаті отримуємо таке рівняння:

$$F = (r^2/k) / (1-r^2) / (n-k-1). \quad (4.10)$$

Розрахунковий F-критерій визначається при відповідному рівні значущості і ступенях свободи, потім його порівнюють з критичним F-критерієм Фішера. Значення останнього критерію представлені в спеціальних таблицях. Якщо розрахунковий F-критерій перевищує його критичне значення, то можна стверджувати, що пояснення, яке дає рівняння в цілому важливе, а економетрична модель адекватна. У протилежному випадку – модель вважається неадекватною, а пояснення неважливе.

Іншим важливим статистичним параметром для перевірки адекватності економетричної моделі є t-розподіл Ст'юдента [70]. Він використовується для оцінки надійності коефіцієнта кореляції. У цьому випадку t-статистика для r розраховується таким чином:

$$t = \sqrt{n-2} / 1-r^2. \quad (4.11)$$

Вибравши дослідним рівень значущості в 5 %, знаходять критичне значення t з $(n-2)$ ступенями свободи. Якщо значення t перевищує його критичне значення (додатній або від'ємний бік), то нульову гіпотезу відхиляють, коефіцієнт кореляції дорівнює нулю. У цьому випадку роблять висновок про лінійний зв'язок (додатній або від'ємний).

Слід зазначити, якщо нульова гіпотеза підтверджується, то значення t буде перевищувати його критичне значення (в додатній або від'ємний бік) тільки в 5 % випадках. Це означає, що при виконанні перевірки ймовірності допущення помилки, що відхиляє нульову гіпотезу, коли вона фактично вірна, складає 5 %.

Ризик допущення такої помилки в 5 % випадків досить великий для дослідника, він може скоротити ступінь ризику, здійснюючи розрахунки при рівні значущості в 1 %. Критичне значення t буде вище, тому необхідна більш

висока (додатна або від'ємна) t-статистика для відхилення нульової гіпотези, а це означає, що потрібне вище значення коефіцієнта кореляції.

Слід вказати і на те, що t-статистика може бути розрахована як співвідношення оцінки коефіцієнта регресії на стандартну помилку.

Розглянемо методику розрахунку F-критерію і t-статистики на прикладі.

Розглянемо приклад. Виконайте відповідні t-тести для багатofакторної моделі. Розрахуйте F-критерій, якщо відомо, що кількість спостережень дорівнює 25, коефіцієнт детермінації (r^2) дорівнює 88 %. Багатofакторна економетрична модель має вигляд:

$$y = 55,3 + 0,093x_1 + 0,087x_2. \quad (4.12)$$

Стандартні помилки дорівнюють:

$$\text{постійний член} - 2,4; x_1 - 0,003; x_2 - 0,002.$$

Розв'язання

Для t-тесту необхідно визначити розрахунковий t-критерій. Для кожного із члена економетричного рівняння він розраховується окремо як співвідношення оцінки коефіцієнта регресії на стандартну помилку.

Таким чином, розрахункові t-критерій такі:

$$t_{p1} = 55,3/2,4 = 23,04;$$

$$t_{p2} = 0,092/0,003 = 30,67;$$

$$t_{p3} = 0,067/0,002 = 43,5.$$

Наступним кроком проведення t-тесту є порівняння розрахункових значень із табличними. Табличне значення t-критерію визначається на основі спеціальних таблиць при відповідних рівнях значущості (5 % або 1 %) і ступенях свободи, які визначаються ($n-k-1$, де n – кількість спостережень; k – кількість факторів моделі, включаючи постійний параметр).

У нашому випадку ступеня свободи дорівнюють:

$$25-4-1 = 20.$$

Табличне значення t-критерію при рівні значущості в 5 % дорівнює 1,725; при 1 % – 2,528.

Як видно, розрахункові значення t-критеріїв всіх факторів моделі значно перевищують його табличні значення. Це означає, що всіх фактори економетричної моделі суттєво впливають на змінний показник (y).

F-критерій визначається за формулою (4.10). Для розробленої економетричної моделі розрахункових F-критеріїв має таке значення:

$$F_p = (0,88/4) / ((1-0,88) / (20)) = 36,7.$$

Потім розрахункове значення F-критерію порівнюємо із його табличним значенням при відповідному рівні значущості і кількості спостережень.

При 5 % рівні значущості для 25 спостережень табличний F-критерій дорівнює 2,99, при 1 % – 29,46.

Таким чином, розрахункові значення F-критерію більші табличних, що вказує на суттєвий рівень пояснення причинно-наслідкових зв'язків економетричної моделі.

Наступним етапом оцінки адекватності економетричної моделі є перевірка її на гетеро- або гомоскедастичність. Гомоскедастичність означає однаковий розподіл фактичних значень вибірки змінних. Тобто фактичні значення спостережень іноді будуть додатними, іноді від'ємними, іноді – відносно близькими до нуля, проте в апriorі відсутні причини появи великих відхилень між спостереженнями.

Разом з тим, для деяких вибірок, більш доцільно припустити, що теоретичний розподіл випадкового члену є різним для різних спостережень. Це не означає, що випадковий член обов'язково буде мати особливо більші (додатні або від'ємні) значення в кінці вибірки, проте це означає, що апriorна ймовірність отримання суттєво відхилених значень буде відносно висока. Це є прикладом гетероскедастичності, що означає «неоднаковий розподіл».

Гетероскедастичність стає проблемою, коли значення змінних, які включаються в рівняння регресії, значно відрізняються в різних спостереженнях. Якщо залежність може бути описана рівнянням, в якому економічні показники змінюють свій масштаб одночасно, то зміна значень невиключених змінних і помилок виміру, впливаючи разом на випадковий член, роблять його порівняно незначними при незначних у і х і порівняно великими – при великих у і х.

Досить часто можна виявити проблему гетероскедастичності. В таких умовах можна здійснити відповідні дії по виключенню цього ефекту на етапі специфікації моделі регресії, і це дозволить зменшити або, можливо, усунути необхідність формальної перевірки. Запропонована значна кількість тестів (і, відповідно, критеріїв для них). Найбільш поширеними тестами є: тест рангової кореляції Спірмена, тест Голдфелда-Квандта і тест Глейзера [67].

При виконанні *тесту рангової кореляції Спірмена* [69] припускається, що дисперсія випадкового члену буде або збільшуватися, або зменшуватися відповідно до збільшення змінної x , і тому в регресії абсолютні значення залишків і значення x будуть корельовані. Дані по x і залишки упорядковуються, коефіцієнт рангової кореляції визначається як:

$$r_{x,e} = 1 - (6 \sum D_i^2 / n(n^2 - 1)), \quad (4.13)$$

де D_i – різниця між рангом x і рангом помилки e ;

e – залишки.

Якщо припускати, що відповідний коефіцієнт кореляції для генеральної сукупності дорівнює нулю, то коефіцієнт рангової кореляції має нормальний розподіл з математичним очікуванням 0 і дисперсією $1/(n-1)$ в більших вибірках. Таким чином, відповідна тестова статистика дорівнює $r_{x,e} \sqrt{n-1}$, і при використанні двобокового критерію нульова гіпотеза про відсутність гетероскедастичності буде відхилена при рівні значущості в 5 %, якщо вона перевищує 1,96, і при рівні значущості в 1 %, якщо вона перевищує 2,58. Якщо в моделі регресії знаходиться більше однієї пояснювальної змінної, то перевірка гіпотези може здійснюватися з використанням іншої з них.

Найбільш відомим формальним критерієм є критерій, запропонований *С. Голдфелдом і Р. Квандтом* [70].

При проведенні перевірки по цьому критерію припускають, що стандартне відхилення (σ_i) розподілу ймовірностей U_i пропорційно значенню x в цьому спостереженні. Запропоновано також, що випадковий член розподілений нормально і не піддається автокореляції.

Всі n спостережень у виборці упорядковуються по значенню x , після чого оцінюються окремі регресії для перших n' і для останніх n' спостережень; середні $(n-2n')$ спостереження відхиляються. Якщо припущення відносно природи гетероскедастичності доцільно, то дисперсія U і в останніх n' спостереженнях буде більшою, чим в перших n' , і це буде відображено в сумі квадратів залишків в двох вказаних «часткових» регресіях [72].

Визначають суми квадратів залишків в регресіях для перших n' і останніх n' спостережень відповідно через RSS_1 і RSS_2 . Розрахуємо відношення RSS_2/RSS_1 , яке має F-розподіл з $(n'-k-1)$ і $(n'-k-1)$ ступенями свободи, де k – число пояснювальних змінних в регресійному рівнянні.

Потужність критерію залежить от вибору n' по відношенню до n . Грунтуючись на результатах деяких проведених експериментів, С. Голдфелд і Р. Квандт стверджують, що n' повинно складати порядок 11, коли $n = 30$, і порядку 22, коли $n = 60$.

Якщо в моделі знаходиться більше однієї пояснювальної змінної, то спостереження повинні упорядковуватися по тій з них, яка пов'язана з σ_i і n' повинно бути більше, ніж $k + 1$ (де k – число пояснювальних змінних).

Метод Голдфелда-Квандта [71] може бути також використаний для перевірки на гетероскедастичність при припущенні, що σ_i обернено пропорційне x_i . При цьому використовується подібна процедура, проте тестова статистика є показником RSS_1 / RSS_2 , який знову має F-розподіл з $(n'-k-1)$ і $(n'-k-1)$ ступенями свободи.

Тест Глейзера [71] дозволяє більш ретельно розглянути характер гетероскедастичності. Він ґрунтується на тому, що знімається припущення, що σ_i пропорційна x_i , а перевіряється лише більш подібна функціональна форма.

Для того, щоб використовувати цей метод, необхідно оцінити регресійну залежність y від x за допомогою методу найменших квадратів, а потім розрахувати абсолютні значення залишків e , оцінивши їх регресію. В кожному випадку нульова гіпотеза про відсутність гетероскедастичності буде відхилена, якщо оцінка регресії відрізняється від нуля. Якщо оцінюється більше однієї функції, то орієнтиром при визначенні характеру гетероскедастичності може бути найкраща з них.

У цьому розділі представлені основні критерії й тести щодо оцінки адекватності моделі. В економетричних дослідженнях можна використовувати й інші тести і критерії. Представлені критерії оцінки адекватності економетричної моделі дають змогу отримати більш ґрунтовні і, насамперед, об'єктивні результати процесів, які відбуваються на підприємстві.

В економетричному моделюванні необхідно враховувати також явище мультиколінеарності.

Мультиколінеарність – це явище, яке використовується для опису проблеми, коли нестрога лінійна залежність між пояснювальними змінними призводить до отримання ненадійних оцінок регресії [70]. Проте, така залежність, зовсім необов'язково дає незадовільні оцінки. Якщо всі інші умови задовільні, тобто якщо кількість спостережень і вибіркві дисперсії

пояснювальних змінних великі, а дисперсія випадкового члену – мала, то в результаті можна отримати досить позитивні оцінки.

Мультиколінеарність виникає за рахунок отримання нестрогої залежності одного (або більше) незадовільних умов, і це питання ступеня визначеності явища, а не його виду. Оцінки регресії будуть незадовільними у відповідному ступені, коли всі незалежні змінні будуть абсолютно некорельовані. Розгляд цієї проблеми починається тільки тоді, коли вона досить суттєво впливає на результати оцінки регресії.

Досить простий спосіб виявлення мультиколінеарності це побудова матриці коефіцієнтів парної кореляції, яка відображає силу зв'язків між факторами. У випадку, коли коефіцієнти парної кореляції між незалежними факторами входять у відповідний проміжок (табл. 4.4), можна свідчити про рівень мультиколінеарності.

Таблиця 4.4 – Рівень мультиколінеарності в залежності від значень коефіцієнтів парної кореляції між незалежними факторами

Значення коефіцієнтів парної кореляції між незалежними факторами	Рівень мультиколінеарності
$r_{x_1x_2} = 0,85 - 1,0$	сильна
$r_{x_1x_2} = 0,55 - 0,84$	помірна
$r_{x_1x_2} = 0,25 - 0,54$	слаба
$r_{x_1x_2} = 0 - 0,24$	відсутня

Слід також зазначити, що явище мультиколінеарності, тобто лінійної залежності одного із аргументу від інших, виявляється декількома способами:

- професійними міркуваннями по суті досліджуваного явища;
- інструкцією на основі «внутрішніх» і «зовнішніх» коефіцієнтів кореляції кожного з аргументів. Якщо «внутрішній» коефіцієнт кореляції більше «зовнішнього», то даний аргумент у рівняння множинної кореляції не слід включати;
- використанням статистичного критерію мультиколінеарності (Феррара і Гюбера). Для цього розглядається величина:

$$\bar{f}_j = (C_{ij} - 1) \frac{n - p}{p - 1}, \quad (4.14)$$

де C_{ij} – діагональні елементи матриці, зворотної до кореляційної, знайденої за вибірковими даними;

n – обсяг вибірки;

p – число аргументів у рівнянні множинної регресії.

Зворотною по відношенню до даної називається матриця, яка, будучи помноженою як справа, так і зліва на дану матрицю, дає одиничну матрицю.

Для матриці A зворотна їй позначається через A^{-1} . Тоді за визначенням маємо:

$$A^{-1} * A = A * A^{-1} = E. \quad (4.15)$$

Якщо існує зворотна матриця A^{-1} , то матриця A називається зворотною. Для виродженої матриці зворотної матриці не існує, оскільки її визначник рівний нулю.

Визначник зворотної матриці рівний зворотній величині визначника даної матриці, що дає можливість обчислення зворотної матриці за допомогою визначників. Для цього використовуються поняття мінору і доповнення алгебри.

Мінором M_{ij} елемента a_{ij} визначника $D = (O_{ij})$ називається такий новий визначник, який отриманий з даного визначника викреслюванням рядка і стовпця, що проходить через даний елемент матриці A .

Доповненням алгебри елемента a_{ij} визначника називається мінор M_{ij} цього елемента, взятий зі знаком $(-1)^{i+j}$. Доповнення алгебри елемента a_{ij} позначається через A_{ij} . У прийнятому нами позначенні матимемо:

$$A^{-1} = \begin{vmatrix} \frac{A_{11}}{D} & \frac{A_{21}}{D} & \dots & \frac{A_{n1}}{D} \\ \frac{A_{21}}{D} & \frac{A_{22}}{D} & \dots & \frac{A_{n2}}{D} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_{1n}}{D} & \frac{A_{2n}}{D} & \dots & \frac{A_{nn}}{D} \end{vmatrix}. \quad (4.16)$$

Ферраром і Глобером доведено, що статична величина \bar{f}_j підпорядковується розподілу Фішера з $(n-p)$ і $(p-1)$ ступенями свободи [72]. Отже, для виявлення мультиколінеарності використовується звичайний прийом перевірки статистичних гіпотез.

Обчисливши вираз

$$\bar{f}_j \quad (j = 1, 2, \dots, p),$$

порівнюємо значення з табличними значеннями $\bar{f}_{5\%}$ і $\bar{f}_{1\%}$ при відповідних ступенях свободи $[(n-p) (p-1)]$.

Якщо

$$\bar{f}_j < \bar{f}_{5\%},$$

то гіпотеза відсутності мультиколінеарності j -го аргументу з іншими в генеральній сукупності стверджується.

Навпаки, при

$$\bar{f}_j > \bar{f}_{5\%}$$

відкидається гіпотеза відсутності мультиколінеарності j -го аргументу з іншими в генеральній сукупності.

При

$$\bar{f}_{5\%} < \bar{f}_j < \bar{f}_{1\%}$$

використовуються засоби послаблення мультиколінеарності шляхом переходу до нелінійних залежностей.

Висновки про виключення якогось аргументу супроводжуються логічним аналізом. По аргументах, що збереглися, повторюється перевірка мультиколінеарності.

Існують різні методи для зменшення мультиколінеарності. Вони діляться на дві категорії: до першої категорії відносяться методи спрямовані на виконання умов, що забезпечують надійність оцінок регресії; до другої – відноситься використання зовнішньої інформації. Якщо використовувати можливі значення показників, то було б важливим збільшити кількість спостережень. Якщо, наприклад, використовуються часові ряди, то це можна зробити шляхом скорочення терміну кожного періоду часу.

Якщо використовуються дані перехресної вибірки і дослідник знаходиться на стадії планування дослідження, то можна збільшити точність оцінок регресії і послабити проблему мультиколінеарності за рахунок більших витрат коштів на збільшення розміру вибірки.

Слід зазначити, що ці методи лише зменшують вплив мультиколінеарності. У практиці економетричного моделювання процесів нівелювання впливу цього явища здійснюється шляхом виключення одного з незалежних факторів моделі, який сильно впливає на інший фактор, а потім продовжують дослідження.

Парний регресійний аналіз спрямований на визначення ступеню зв'язку між змінними, і яким чином вони пов'язані в побудові парної моделі [70]. Слід зазначити, що не слід очікувати отримання точного співвідношення між будь-якими економічними показниками, крім випадків, коли воно існує по визначенню.

Парний регресійний аналіз відбувається за *такими напрямками* [70]:

1. Збір статистичної інформації, яка відображає економічні процеси на підприємстві. Цей процес відбувається шляхом обробки фінансових, економічних, бухгалтерських, статистичних документів діяльності суб'єктів підприємницької діяльності.

2. Обробка статистичної інформації, її специфікація. Це важливий етап, оскільки він створює підґрунтя для отримання об'єктивних результатів і адекватної парної економетричної моделі.

3. Використання економетричного інструментарію для розробки парної моделі. В цьому аспекті здійснюється побудова матриці статистики, оцінка показників варіації змінних, розрахунок коефіцієнтів парної кореляції й детермінації, визначення показників параметрів парної економетричної моделі. Слід зазначити, що для обчислення параметрів рівняння виду:

$$\bar{y} = kx + b$$

(лінійна парна модель залежності) частіш за все користуються методом найменших квадратів. При цьому ставиться умова, щоб сума квадратів відхилень (відстаней) всіх досліджених точок від ординат, обчислених за рівнянням прямої ε_i , була мінімальною. Іншими словами, пряма повинна проходити якомога ближче до вершин емпіричної лінії регресії. Це означає, що параметри k і b управління регресії необхідно визначити з рівняння:

$$\varepsilon_i = \sum_{i=1}^n (y_i - \tilde{y}_i) = \min, \quad (4.17)$$

де y_i – ординати досліджуваних точок;

\tilde{y}_i – ординати розрахункових точок, визначені за рівнянням регресії:

$$\bar{y}_i = kx_i + b,$$

таким чином,

$$\varepsilon_i = \sum_{i=1}^n [y_i - (kx + b)]^2 = F(k, b) \min.$$

Умовою екстремуму даної функції слід вважати рівність нулю часткових похідних, узятих за параметрами k і b :

$$\frac{dF}{dk} = 0,$$

$$\frac{dF}{db} = 0,$$

$$[F(u)]' = F_u(u) *',$$

звідси,

$$\begin{aligned}\frac{dF}{dk} &= -2 \sum_{i=1}^n [y_i - (kx_i + b)] x_i = 0, \\ \frac{dF}{db} &= -2 \sum_{i=1}^n [y_i - (kx_i + b)] = 0.\end{aligned}\tag{4.18}$$

Скоротивши на (-2) і розкривши квадратні дужки, отримаємо систему лінійних рівнянь:

$$\begin{aligned}\sum_{i=1}^n y_i x_i &= k \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i, \\ \sum_{i=1}^n y_i x_i &= k \sum_{i=1}^n x_i + bn,\end{aligned}\tag{4.19}$$

підставивши сюди чисельні значення відповідних величин, знайдемо параметри k і b .

У разі лінійної залежності геометричне і алгебраїчне значення коефіцієнта регресії полягає в тому, що він кількісно характеризує на скільки в середньому змінюється y при зміні x_i на одиницю свого вимірювання. Чим більші чисельні значення коефіцієнта регресії, тим більший відносний приріст функції при зміні аргументу.

4. Оцінка адекватності розробленої парної лінійної економетричної моделі на основі відповідних критеріїв і тестів.

5. Інтерпретація отриманих параметрів парної лінійної економетричної моделі. Це важливий етап, на якому відображається економічна результативність економетричного моделювання.

У загальному вигляді економетрична модель парної лінійної регресії може мати вигляд:

$$y = a_0 + a_1x + e, \quad (4.20)$$

де y – залежна змінна;

x – незалежна змінна;

a_0, a_1 – параметри економетричного рівняння;

e – випадковий член.

Таким чином, парний регресійний аналіз дозволяє побудувати парну лінійну економетричну модель, яка дозволить встановити причинно-наслідковий зв'язок між залежною економічною змінною і незалежним фактором і створити передумови для побудови організаційно-економічних механізмів управління підприємствами та прийняття рішень, спрямованих на розвиток цих суб'єктів підприємницької діяльності. Проте, більшість економічних процесів мають складний характер, де враховується велика кількість факторів. Тому необхідно будувати економетричні моделі, які враховують декілька економічних показників, тобто розробляти лінійні моделі множинної регресії.

4.2 Лінійні моделі множинної регресії

Кількісний регресійний аналіз є продовженням парного регресійного аналізу у випадках, коли залежна змінна y пов'язана з двома або більше незалежними змінними x . Тобто відбувається розширення парної регресійної моделі, де важливе значення відіграють спільний вплив незалежних змінних на залежну змінну. Тому в кількісному регресійному аналізі необхідно враховувати й чітко визначати цей вплив, а також важливе значення має вирішення проблеми специфікації.

Остання проблема лежить в площині вибору тих факторів, які впливають на результуючий показник, економічно інтерпретуються і об'єктивно відображають господарські процеси, що відбуваються на підприємстві.

Результатом кількісного регресійного аналізу є побудова кількісної (багатофакторної) регресійної моделі.

У загальному вигляді, кількісна регресійна модель має вигляд:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_ix_i + e$$

або

$$y = b + k_1x_1 + k_2x_2 + \dots + k_ix_i + e, \quad (4.21)$$

де y – результуюча залежна змінна;

x_1, x_2, x_i – незалежна змінна;

$a_0, a_1, a_2, a_i, b, k_1, k_2, k_i$ – параметри рівняння (коефіцієнти регресії);

e – випадковий член.

У кількісному регресійному аналізі визначається *коефіцієнт регресії*, який необхідний для забезпечення найкращої відповідності спостереженням і отримання оптимальних оцінок невідомих значень параметрів моделей.

Для *розрахунку коефіцієнтів регресії* $a_0, a_1, a_2, \dots, a_i$ використовують метод найменших квадратів. Так для пошуку коефіцієнтів регресії (параметрів) двофакторної моделі складають систему нормальних рівнянь:

$$\left. \begin{aligned} \sum y &= n \cdot a_0 + a_1 \sum x_1 + a_2 \sum x_2 \\ \sum yx_1 &= a_0 \sum x_1 + a_1 \sum x_1^2 + a_2 \sum x_1x_2 \\ \sum yx_2 &= a_0 \sum x_2 + a_1 \sum x_1x_2 + a_2 \sum x_2^2 \end{aligned} \right\}. \quad (4.22)$$

Кількісний регресійний аналіз дозволяє розмежовувати вплив незалежних змінних, допускаючи при цьому можливість їх корелювати. Коефіцієнт регресії для кожної змінної x дає оцінку її впливу на величину y в випадку незмінності впливу на неї всіх інших змінних x .

Це може бути встановлено двома способами. Один з них складається в виявленні того, що, якщо модель правильно специфікована і виконуються умови Гауса-Маркова, то оцінки будуть незміщеними. Інший спосіб полягає в оцінюванні регресійної залежності y від однієї з незалежних змінних, усуненні перед цим можливості використання останньої як заміщуючої для інших будь-яких незалежних змінних і показавши далі, що оцінка її коефіцієнта регресії співпадає з оцінкою коефіцієнта кількісної регресії. У рамках висвітлених способів необхідно розглянути умови Гауса-Маркова [13].

Якість коефіцієнтів регресії залежить від якості випадкового члена. Для того, щоб регресійний аналіз давав найкращі результати, випадковий член повинен задовольняти 4 умовам, відомими як умови Гауса-Маркова [13].

1-а умова Гауса-Маркова полягає в тому, що математичне очікування випадкового члена будь-якого спостереження повинно дорівнювати нулю.

2-а умова Гауса-Маркова полягає в тому, що дисперсія випадкового члена повинна бути постійною для всіх спостережень.

3-а умова Гауса-Маркова припускає відсутність систематичного зв'язку між значення випадкового члену в будь-яких спостереженнях. Випадкові члени повинні бути абсолютно незалежними один від одного.

4-а умова Гауса-Маркова полягає в тому, що випадковий член повинен бути розподілений незалежно від пояснювальних змінних. Тобто пояснювальні змінні не є стохастичними. Значення будь-якої незалежної змінної в кожному спостереженні повинно бути визначено зовнішніми причинами, які не визначені в рівнянні регресії.

Коефіцієнти регресії є більш точними [69]:

- чим більша кількість спостережень у виборці;
- чим більша дисперсія вибірки пояснювальних змінних;
- чим менша теоретична дисперсія випадкового члену;
- чим менше зв'язані між собою пояснювальні змінні.

Стандартна помилка коефіцієнта кількісної регресії визначається аналогічно, як і в парному регресійному аналізі. Тобто формула для стандартної помилки може бути визначена на основі заміни дисперсії на незміщену оцінку і отримання квадратного кореню.

Результатом кількісного регресійного аналізу є побудова багатofакторної моделі, яка відображає причинно-наслідкові зв'язки між факторами і створює кількісне підґрунтя для розробки механізмів і прийняття ефективних управлінських рішень.

Для побудови лінійної моделі множинної регресії використовується статистична інформація про діяльність підприємства і здійснюються такі етапи: математико-статистичний аналіз, побудова багатofакторної регресійної моделі, перевірка побудованої моделі на адекватність, аналіз (інтерпретація) отриманих результатів.

На етапі математико-статистичного аналізу проводиться перевірка основних припущень класичного регресійного аналізу, крім того, здійснюється найважливіша процедура багатofакторного аналізу – перевірка факторів на мультиколінеарність [43].

Слід зазначити, що термін «мультиколінеарність» означає, що в багатofакторній регресійній моделі дві або більше незалежних змінних (факторів) пов'язані між собою лінійною залежністю або, іншими словами, мають високий ступінь кореляції ($r_{x_i} \rightarrow 1, i \neq j$).

Для здійснення математико-статистичного аналізу будується матриця коефіцієнтів парної кореляції, який показує ступінь зв'язку між факторами економетричної моделі. Потім аналізуються коефіцієнти парної кореляції між факторами. Результатом етапу математико-статистичного аналізу є знаходження множини основних незалежних між собою факторів, що є базою для побудови регресійної моделі.

На другому етапі для побудови багатofакторної моделі вибираються фактори, що будуть відображати причинно-наслідковий зв'язок. В цьому аспекті широке використання отримали «покроковий» метод і метод «виключень».

Найбільш доцільно відшукувати рівняння множинної регресії шляхом послідовного підключення до парного рівняння решти аргументів в порядку їх значущості («покроковий метод»).

У цьому випадку виявляється можливість на кожному етапі аналізувати:

- обумовленість вирішуваної системи за чисельним значенням її визначника (детермінатора);
- зміну β - коефіцієнтів, чисельне значення яких має бути менше 1, а знак не суперечити логіці;
- зростання коефіцієнта множинної кореляції R і убування залишкової дисперсії $\tilde{\sigma}_{oct}^2 = \tilde{\sigma}_y^2(1 - R^2)$.

Методика послідовного підключення аргументів складається з таких операцій [66]:

1. Обирається аргумент x_l , якому відповідає найбільший за абсолютним значенням «зовнішній» коефіцієнт кореляції:

$$|r_{y_l}| = \max |r_{y_j}|, j = 1, 2, \dots, q. \quad (4.23)$$

Як аргумент x_l записується рівняння:

$$t_{y_l} = t_{y_l} t_{x_l}. \quad (4.24)$$

2. Приєднується аргумент x_{j_0} , для якого:

$$|r_{x_j} x_1| = \min |r_{x_j} x_1|, j = 2, 3, \dots, q. \quad (4.25)$$

Складається система нормальних рівнянь:

$$r_{yx_1} = \beta_1 + r_{x_{j_0}} \beta_2, \quad (4.26)$$

$$r_{yx_{j_0}} = \beta_1 r_{x_{j_0} x_1} + \beta_2, \quad (4.27)$$

і обчислюються значення β_1 та β_2 .

Визначаються:

$$R^2_{y, x_1, x_{j_0}} = \beta_1 r_{yx_1} + \beta_2 r_{yx_{j_0}}, \quad (4.28)$$

$$\sigma_{y, x_1, x_{j_0}} = \sqrt{1 - R^2_{y, x_1, x_{j_0}}}. \quad (4.29)$$

Порівнюється $R^2_{y, x_1, x_{j_0}}$, $\sigma_{y, x_1, x_{j_0}}$ відповідно до $r^2_{yx_1}$, σ_{yx_1} .

Переконаються в справедливості нерівності:

$$R^2_{y, x_1, x_{j_0}} \geq r^2_{yx_1}; \sigma_{y, x_{j_0}} \leq \sigma_{y, x_1}. \quad (4.30)$$

У протилежному разі замінюється аргумент іншим x_{j_l} , а аргумент x_{j_0} переноситься на останнє місце.

3. Далі приєднується наступний аргумент x_{j_l} , і розв'язується система з трьома невідомими:

$$r_{yx_1} = \beta_1 + \beta_2 r_{x_1 x_{j_0}} + \beta_3 r_{x_1 x_{j_l}}, \quad (4.31)$$

$$r_{yx_{j_0}} = \beta_1 r_{x_1 x_{j_0}} + \beta_2 + \beta_3 r_{x_{j_0} x_{j_l}}, \quad (4.32)$$

$$r_{yx_{j_l}} = \beta_1 r_{x_1 x_{j_0}} + \beta_2 r_{x_{j_0} x_{j_l}} + \beta_3. \quad (4.33)$$

Обчислюються значення β_1 , β_2 та β_3 .

Визначаються:

$$R^2_{y, x_1, x_{j_0}, x_{j_l}} = \beta_1 r_{yx_1} + \beta_2 r_{yx_{j_0}} + \beta_3 r_{yx_{j_l}}, \quad (4.34)$$

$$\sigma_{y, x_{j_0}, x_{j_1}} = \sigma_y \sqrt{1 - R^2_{y, x_1, x_{j_0}, x_{j_1}}} \quad (4.35)$$

і порівнюються з $R^2_{y, x_1, x_{j_0}}$ і $\sigma_{y, x_1, x_{j_0}}$.

Переконаються в справедливості нерівностей:

$$R^2_{y, x_1, x_{j_0}, x_{j_1}} \geq R^2_{y, x_1, x_{j_0}}, \quad (4.36)$$

$$\sigma_{y, x_1, x_{j_0}, x_{j_1}} \leq \sigma_{y, x_1, x_{j_0}}. \quad (4.37)$$

У протилежному разі поступають аналогічно пункту 2.

Дослідження ведуть до тих пір, поки не будуть апробовані чинники-аргументи і збережені тільки ті з них, для яких β_j коефіцієнти суттєві й лінійно незалежні. У результаті виходить множинне рівняння в стандартизованому масштабі.

Від рівняння множинної регресії в стандартизованому масштабі

$$t_{\bar{y}_{x_i}} = \beta_1 t_1 + \beta_2 t_2 + \beta_p t_n \quad (4.38)$$

до рівняння множинної регресії в натуральному масштабі:

$$\bar{y}_{x_1, x_2, \dots, x_p} = a_1 x_1 + a_2 x_2 + \dots + a_p x_p + b. \quad (4.39)$$

Перехід здійснюється подвійно. Шляхом використання формули

$$t_{\bar{y}_{x_i}} = \frac{y - \bar{y}}{\sigma_y}, \quad t_{x_i} = \frac{x_i - \bar{x}_i}{\sigma_{x_i}}, \quad (4.40)$$

де $i = 1, 2, \dots, p$.

При цьому маємо

$$\frac{y - \bar{y}}{\sigma_y} = \beta_1 \frac{x_1 - \bar{x}_1}{\sigma_{x_1}} + \beta_2 \frac{x_2 - \bar{x}_2}{\sigma_{x_2}} + \dots + \beta_p \frac{x_p - \bar{x}_p}{\sigma_{x_p}}. \quad (4.41)$$

Підставивши відомі значення \bar{y} , σ_{x_i} , σ_y , β_i і \bar{x}_1 , отримаємо рівняння множинної регресії в натуральному масштабі, в якому чисельне значення вільного члена додатково визначати непотрібно.

2. Невідомі коефіцієнти a_i в рівнянні множинної регресії в натуральному масштабі визначають з виразу

$$a_1 = \beta_1 \frac{\sigma_y}{\sigma_{x_1}}, a_2 = \beta_2 \frac{\sigma_y}{\sigma_{x_2}}, \dots, a_p = \beta_p \frac{\sigma_y}{\sigma_{x_p}}. \quad (4.42)$$

Чисельне значення вільного члена:

$$b = \bar{y} - (a_1 \bar{x}_1 + a_2 \bar{x}_2 + \dots + a_p \bar{x}_p). \quad (4.43)$$

Метод «виключень» полягає в тому, що вибирається набір факторів, які ймовірно можуть впливати на результативний показник. Потім по черзі виключаються ті фактори, у який найменший коефіцієнт кореляції (згідно матриці статистики), а значення часткових F-критеріїв не перевищують нормативні значення. Таким чином, залишаються лише ті змінні, які відповідають розглянутим вище умовам.

Слід вказати, що на цьому етапі розраховується *коефіцієнт множинної кореляції*, який показує загальний вплив незалежних факторів на результуючий показник економетричної моделі. Він знаходиться у проміжку між 0 і 1.

Чим більший вплив факторів, тим коефіцієнт множинної кореляції наближається до 1. Він не може перевищувати значення останньої.

Розрахунок коефіцієнта множинної кореляції ($R_{yx_1x_2\dots x_n}$) розраховується за формулою Боярського [67]:

$$R_{yx_1x_2\dots x_n} = \sqrt{\frac{-1^\alpha \Delta_*}{\Delta_0}}, \quad (4.44)$$

де α – порядок повної матриці коефіцієнтів кореляції;

Δ_* – визначник повної матриці коефіцієнтів кореляції із заміною нижнього правого елементу нулем;

Δ_0 – визначник матриці, в якій враховані коефіцієнти парної кореляції незалежних факторів.

Якщо розкрити визначники для двофакторної економетричної моделі, то коефіцієнт множинної кореляції може бути визначений:

$$R_{yx_1x_2} = \sqrt{\frac{\begin{vmatrix} 1 & r_{x_1x_2} & r_{yx_1} \\ r_{x_1x_2} & 1 & r_{yx_2} \\ r_{yx_1} & r_{yx_2} & 0 \end{vmatrix}}{\begin{vmatrix} 1 & r_{x_1x_2} \\ r_{x_1x_2} & 1 \end{vmatrix}}}, \quad (4.45)$$

де r_{yx_1} , r_{yx_2} – коефіцієнти парної кореляції між залежною змінною y і незалежними факторами x_1 , x_2 ;

$r_{x_1x_2}$ – коефіцієнт парної кореляції між незалежними змінним x_1 , x_2 .

З метою контролю правильності розрахунків цей коефіцієнт визначають також за формулою [67]:

$$R_{yx_1x_2\dots x_n} = \sqrt{r_{yx_1}\beta_{x_1} + r_{yx_2}\beta_{x_2} + \dots + r_{yx_n}\beta_{x_n}}, \quad (4.46)$$

де $\beta_{x_1,x_2,\dots,x_n}$ – β -коефіцієнти для незалежних факторів економетричної моделі.

Цей коефіцієнт може бути розрахований таким чином [67]:

$$\beta_i = \Delta_i / \Delta_0, \quad (4.47)$$

де Δ_i – визначник (детермінант) матриці взаємної кореляції (мультиколінеарності) із заміною в ній i -го стовпця стовпцем коефіцієнтів кореляції r_{yx_i} .

Наприклад, β -коефіцієнти для одного з факторів двофакторної моделі розраховуються таким чином:

$$\beta_{x_1} = \frac{\begin{vmatrix} r_{yx_1} & r_{x_1x_2} \\ r_{yx_2} & 1 \end{vmatrix}}{\begin{vmatrix} 1 & r_{x_1x_2} \\ r_{x_1x_2} & 1 \end{vmatrix}}. \quad (4.48)$$

Знайдені в результаті рішення кореляційної матриці β -коефіцієнти показують на яку частину середньоквадратичного відхилення σ_y змінюється середнє значення функції, якщо відповідний аргумент зменшується або збільшується, а інші аргументи залишаються незмінними.

Для з'ясування математико-статистичного змісту множинної кореляції всю досліджувану групу змінних слід розглядати як один чинник-аргумент. При цьому розраховується коефіцієнт надійності:

$$M = \frac{R\sqrt{n}}{1 - R^2}. \quad (4.49)$$

Стандартну помилку (середню квадратичну похибку) коефіцієнта множинної кореляції визначають за формулою:

$$\sigma_R = (1 - R) / \sqrt{n}, \quad (4.50)$$

де n – обсяг вибірки.

Сукупний вплив врахованих змінних на функцію визначається коефіцієнтом загальної детермінації R^2 , а окремих чинників-аргументів за чисельними значеннями детермінації $r_i\beta_i$:

$$R^2 = r_1\beta_1 + r_2\beta_2 + \dots + r_p\beta_p. \quad (4.51)$$

Стандартну (систематичну) похибку \hat{R}^2 обчислюють за формулою

$$\hat{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-p}, \quad (4.52)$$

де p – число параметрів рівняння регресії.

З рівняння множинної регресії можна отримати рівняння чистої регресії у по кожному з аргументів x_i . Для цього фіксується значення всіх аргументів, окрім x_i , на середньому рівні.

Отримане рівняння описує, як в середньому змінюється змінна x_i , якщо всі інші аргументи постійні й закріплені саме на своїх середніх рівнях.

Розглянемо приклад. Розрахуйте коефіцієнт множинної кореляції та визначте β -коефіцієнти, на основі даних представлених в таблиці 4.5.

Таблиця 4.5 – Матриця статистики економічних показників

Показники	Коефіцієнти парної кореляції		
	$P(y)$ (рентабельність продукції)	$\Phi\text{Зоз}(x_1)$ (фондоозброєність основних засобів)	$Ч(x_2)$ (середньоспискова чисельність працівників)
$P(y)$ (рентабельність продукції)	1	0,87	0,65
$\Phi\text{Зоз}(x_1)$ (фондоозброєність основних засобів)	0,87	1	0,36
$Ч(x_2)$ (середньоспискова чисельність працівників)	0,65	0,36	1

Розв'язання

Визначимо β - коефіцієнти для факторів x_1 і x_2 , формула (4.48):

$$\beta_{x_1} = \frac{\begin{vmatrix} 0,87 & 0,36 \\ 0,65 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 0,36 \\ 0,36 & 1 \end{vmatrix}} = 0,731,$$

$$\beta_{x_2} = \frac{\begin{vmatrix} 0,87 & 1 \\ 0,65 & 0,36 \end{vmatrix}}{\begin{vmatrix} 1 & 0,36 \\ 0,36 & 1 \end{vmatrix}} = 0,387.$$

Розрахуємо коефіцієнт множинної кореляції, формула (4.46):

$$R_{yx_1x_2} = \sqrt{r_{yx_1}\beta_{x_1} + r_{yx_2}\beta_{x_2}} = \sqrt{0,87 \cdot 0,7371 + 0,65 \cdot 0,387} = 0,945.$$

На наступному етапі аналізу перевіряється адекватність моделі за допомогою використанням F-критерію Фішера і t-критерію Ст'юдента.

При перевірці на адекватність економетричної моделі також використовується тест Дарбіна-Уотсона, який спрямований для перевірки кореляції між залишками.

На останньому етапі отримана модель аналізується і інтерпретується.

Статистична оцінка надійності коефіцієнта регресії здійснюється за допомогою t-критерію Ст'юдента. Він застосовується для оцінки тісноти зв'язку між незалежною змінною x і залежною y . При використанні цього критерію формулюється нульова гіпотеза. Потім отримане значення t-розподілу Ст'юдента порівнюється з критичним. Якщо фактичне значення t-розподілу Ст'юдента перевищує критичне, то спростовується нульова гіпотеза, зв'язок між змінними x і y вважається щільним. Якщо ні, то приймається нульова гіпотеза, а фактори моделі вважаються статистично неадекватними і виключаються з моделі при встановленому рівні значущості в 5 % і 1 %.

F-тест використовується для оцінки пояснення, яке дає рівняння в цілому. Якщо фактичне значення F-критерію вище нормативного, то модель адекватна, а її фактори залишаються у рівнянні.

Для перевірки адекватності економетричної моделі використовується *тест Дарбіна-Уотсона*, який спрямований для перевірки кореляції між залишками. Він включає такі етапи:

1. Розраховуються d-статистики для аналізованої вибірки даних. Значення d-статистики лежать у межах від 0 до 4. Показник Дарбіна-Уотсона розраховується таким чином:

$$DW = \frac{\sum_{j=2}^n (e_j - e_{j-1})^2}{\sum_{j=1}^n e_j^2}, \quad (4.53)$$

де e_j – залишки j -го ряду вибірки даних;

e_{j-1} – залишки попереднього j -го ряду вибірки даних.

2. Порівнюються отримані d-статистики з табличними d-статистиками при рівні значущості $\alpha = 0,05$, кількості факторів k , що присутні в моделі, і кількості спостережень n . Якщо розраховане значення d-статистики знаходиться в проміжку від 0 до d_L ($0 < d < d_L$), то це свідчить про наявність позитивної автокореляції.

Якщо значення d потрапляє в зону невизначеності, тобто набуває значення $d_L \leq d \leq d_U$, або $4 - d_U \leq d \leq 4 - d_L$, то не можемо зробити висновок про наявність чи відсутність автокореляції.

Якщо $4 - d_L < d < 4$, то маємо негативну автокореляцію.

Якщо $d_U < d < 4 - d_U$, то автокореляції немає.

Для оцінки адекватності лінійної моделі множинної регресії важливе значення має перевірка її на *гомо- або гетероскедастичність* [72]. Суть цього явища полягає в тому, що варіація кожної ε_i навколо її математичного сподівання не залежить від значення x . Дисперсія кожної ε_i зберігається сталою незалежно від малих чи великих значень факторів, σ_ε^2 не є функцією x_{ij} , тобто:

$$\sigma_\varepsilon^2 \neq f(x_{1i}, x_{2i}, \dots, x_{pi}).$$

Якщо σ_ε^2 не є сталою, а її значення залежать від значень x_{ij} , можемо записати:

$$\sigma_\varepsilon^2 = f(x_{1i}, x_{2i}, \dots, x_{pi}).$$

У цьому разі маємо справу з гетероскедастичністю. Оцінка моделі на наявність гетероскедастичності полягає в тому, що на першому етапі здійснюється тестування моделі на наявність гетероскедастичності, якщо підтверджується гіпотеза про її наявність, то на другому етапі модель виключається.

Тестування лінійної моделі множинної регресії, як і випадку лінійної моделі парної регресії, на гетероскедастичність здійснюється на підставі тесту рангової кореляції Спірмена. Значущість отриманого коефіцієнта рангової кореляції Спірмена перевіряється за допомогою t-критерію Ст'юдента при (n-2) кількості ступенів свободи.

Фактичне значення t-критерію Ст'юдента зіставляється з $t_{кр}$ [70]:

- якщо $t_\phi > t_{кр}$, то підтверджується гіпотеза про наявність гетероскедастичності;

- якщо $t_\phi < t_{кр}$, то приймається гіпотеза про гомоскедастичність.

На *етапі аналізу отриманих результатів* здійснюється інтерпретація отриманої економетричної моделі. На цьому етапі обґрунтовується доцільність отриманих результатів.

Розглянемо, наприклад, економічний зміст моделі залежності суми капіталу і середньооблікової чисельності працівників ($Ч$), співвідношення власного і позикового капіталів ($\frac{BK}{ПК}$), відношення витрат інвестованого капіталу на оплату праці й матеріали ($\frac{Воп}{Вм}$):

$$K_{nop} = -2355,39 + 28,7Ч + 1795,24 \frac{BK}{ПК} + 8,95 \frac{Воп}{Вм}. \quad (4.54)$$

Економетрична багатофакторна модель (4.48) показує, що 82 % коливань нового капіталу (коефіцієнт детермінації – 82 %) обумовлюється трьома факторами:

- середньообліковою чисельністю працівників;
- співвідношенням власного й позикового капіталу;
- відношенням витрат інвестованого капіталу на оплату праці й матеріали.

Статистичні характеристики моделі адекватні.

Фактичні значення t-статистик більші, ніж критичні (табл. 4.6), фактичне значення критерію Фішера:

$$F_{\phi} = 69 > F_{0,05;24} = 3,01$$

також значно перевищує його критичне (табличне) значення.

Значення критерію Дарбіна-Уотсона свідчить про відсутність автокореляції залишків:

$$1,65 < d_{\phi} = 2,14 < 2,35.$$

Величина критерію Спірмена

$$r_s = 0,124$$

свідчить про гомоскедастичність, оскільки отримане значення t-статистики нижче його критичного значення

$$t_{\phi} = 0,628 < t_{кр} = 1,706.$$

Мультиколінеарність між незалежними факторами низька, оскільки коефіцієнти парної кореляції між цими показниками мають значення в проміжку від 0,15 до 0,45.

Таблиця 4.6 – Значення t-статистик для параметрів моделі (4.54)

Параметри	Розрахункові	Критичні
1	2	3
Постійний параметр	2,912	1,711
Середньооблікова чисельність працівників (Ч)	5,278	1,711

Продовження таблиці 4.6

1	2	3
Співвідношення власного й позикового капіталу ($\frac{BK}{PK}$)	2,329	1,711
Відношення витрат інвестованого капіталу на оплату праці й матеріали ($\frac{Воп}{Вм}$)	1,981	1,711

Економічна інтерпретація моделі (4.54) полягає в тому, що між середньообліковою чисельністю працівників і новим капіталом зв'язок лінійний. Зростання середньооблікової чисельності на одного працівника призведе до збільшення обсягу нового капіталу на 28,7 тис. грн. Між новим капіталом та чинниками: співвідношенням власного й позикового капіталу, відношенням витрат інвестованого капіталу на оплату праці й матеріали, також існує лінійний зв'язок.

Збільшення співвідношення власного та позикового капіталу на 10 $\frac{коп.}{грн.}$ призведе до зростання обсягів нового капіталу на 179,52 грн.

Збільшення відношення витрат інвестованого капіталу на оплату праці та матеріали на 10 $\frac{коп.}{грн.}$ призведе до зростання нового капіталу на 0,9 грн.

Слід зазначити, що коли розглянуті фактори моделі (4.54) будуть дорівнювати 0, то новий (отриманий) капітал буде мати значення – 235,54 грн.

Тобто на цей показник у відповідних умовах негативний вплив здійснюють інші фактори. Тому в подальших дослідженнях необхідно враховувати і їх вплив.

Таким чином, кількісний регресійний аналіз дозволяє встановити причинно-наслідковий зв'язок між декількома економічними факторами, побудувати багатофакторні економетричні моделі й розробити ефективні механізми управління підприємством.

4.3 Узагальнені економетричні моделі в оцінці нерухомого майна

Узагальнена економетрична модель – це окрема функція чи система функцій (рівнянь), що описує кореляційно-регресійний зв'язок між показниками, один чи декілька з яких є залежною змінною, а усі інші – незалежними [67].

Узагальнені економетричні моделі є окремим класом математичних моделей і характеризуються такими особливостями [69]:

- економетричними моделями є *моделі прикладні (емпіричні)*;
- економетричними моделями є *моделі дескриптивні*;
- економетричними моделями є *моделі стохастичні*.

Узагальнена економетрична модель у вигляді однієї функції (рівняння) має такий вигляд [70]:

$$y = f(x_1, x_2, \dots, x_n, \varepsilon), \quad (4.55)$$

де y – залежна змінна;

x_1, x_2, \dots, x_n – незалежні змінні;

ε – випадковий член (складова).

Узагальнені економетричні моделі можуть бути системою функцій, які мають такий вигляд [70]:

$$y_s = f_s(x_{s_1}, x_{s_2}, \dots, x_{s_m}, \varepsilon_s), (s = \overline{1, k}), \quad (4.56)$$

де k – кількість рівнянь.

Прикладом такої моделі може бути модель формування доходу Кейнса:

$$\begin{cases} C_t = \beta_0 + \beta_1 y_t + \varepsilon_t, \\ y_t = C_t + I_t, \end{cases} \quad (4.57)$$

де C_t – сукупне споживання;

y_t – національний дохід;

I_t – інвестиції;

β_0, β_1 – параметри моделі.

Інформаційною базою для побудови узагальнених економетричних моделей є статистичні вибірки. Особливістю цих статистичних вибірок є те, що

в економетричних дослідженнях необхідно враховувати кількість спостережень на один фактор повинна перевищувати 16.

До статистичних вибірок, які враховуються при побудові узагальнених економетричних моделей, пред'являються такі вимоги [72]:

- однорідність спостережень (якісна і кількісна);
- точність.

Узагальнені економетричні моделі можуть бути лінійні або нелінійні.

Узагальнена лінійна економетрична модель – це регресійна модель, яка встановлює лінійну залежність між економічними показниками, один з яких є залежною (пояснюваною) змінною, а всі інші – незалежними (пояснюючими) змінними моделі [71].

Залежна змінна для такої моделі розглядається, як ендогенна змінна, а незалежні змінні – як екзогенні.

Теоретична узагальнена лінійна економетрична модель може бути специфікована у такій формі [70]:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_mx_m + \varepsilon, \quad (4.58)$$

де y – залежна (пояснювана) змінна моделі;

x_1, x_2, \dots, x_m – незалежні (пояснюючі) змінні моделі або фактори;

a_0, a_1, \dots, a_m – параметри моделі;

ε – випадковий член;

m – кількість пояснюючих змінних моделі.

Представлена узагальнена модель (4.58) дійсна для всієї генеральної сукупності спостережень за змінними моделі й відображає відповідну економічну ситуацію, яка склалась на макро- або мікрорівні.

Вибіркова узагальнена лінійна економетрична модель має такий вигляд:

$$y = A_0 + A_1x_1 + A_2x_2 + \dots + A_mx_m + e, \quad (4.59)$$

де y – залежна (пояснювана) змінна моделі;

x_1, x_2, \dots, x_m – незалежні (пояснюючі) змінні моделі (фактори);

A_0, A_1, \dots, A_m – параметри вибіркової моделі;

e – залишки моделі;

m – кількість пояснюючих змінних моделі.

Вибіркова модель (4.59) розробляється для певної статистичної вибірки з генеральної сукупності. На відміну від моделі (4.58) параметри вибіркової

моделі A_0, A_1, \dots, A_m є оцінками (наближеними значеннями) параметрів a_0, a_1, \dots, a_m і випадковими величинами, а залишки моделі e можна оцінити на основі статистичних даних.

Таким чином, вибіркова модель завжди є тільки оцінкою реальної, але невідомої теоретичної моделі.

Вибіркова функція регресії для узагальненої лінійної економетричної моделі має такий вигляд [70]:

$$\hat{y} = A_0 + A_1 x_1 + A_2 x_2 + \dots + A_m x_m, \quad (4.60)$$

де \hat{y} – оцінка математичного сподівання залежної (пояснюваної) змінної моделі;

x_1, x_2, \dots, x_m – незалежні (пояснюючі) змінні моделі (фактори);

A_0, A_1, \dots, A_m – параметри вибіркової регресії;

m – кількість пояснюючих змінних моделі.

Узагальнена нелінійна економетрична модель – це регресійна модель, яка встановлює нелінійну залежність між показниками.

Узагальненими нелінійними економетричними моделями можуть бути відомі функції [70]:

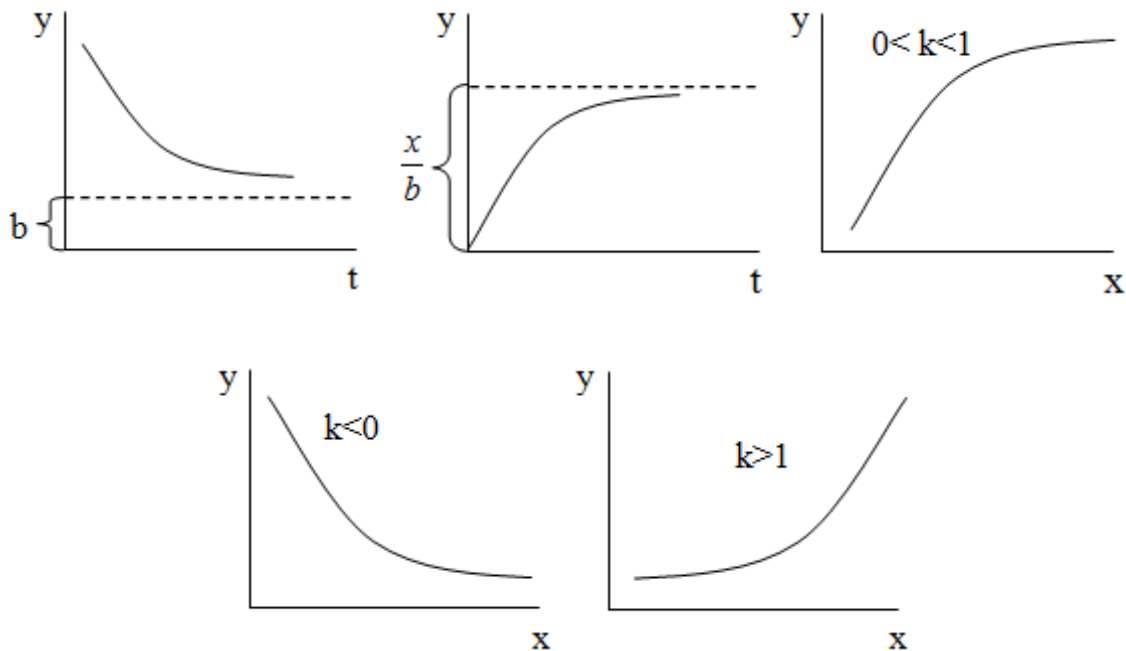
1. Квадратична – $\hat{y} = a_0 + a_1 x + a_2 x^2$ або $\hat{y} = b + k_1 x + k_2 x^2$.
2. Гіперболічна – $\hat{y} = a_0 + \frac{a_1}{x}$ або $\hat{y} = b + \frac{k}{x}$.
3. Степенева – $\hat{y} = a_0 x^{a_1}$ або $\hat{y} = b x^k$.
4. Модифікована експонента – $\hat{y} = a_0 + a_1 a_2^x$ або $\hat{y} = b + k_1 k_2^x$.
5. Крива Гомперця – $\hat{y} = e^{a_0 + a_1 a_2^x}$ або $\hat{y} = e^{b + k_1 k_2^x}$.
6. Логістична – $\hat{y} = 1/(a_0 + a_1 a_2^x)$ або $\hat{y} = 1/(b + k_1 k_2^x)$.
7. Показова функція – $\hat{y} = a_0 e^{a_1 x}$ або $\hat{y} = b e^{kx}$.

Більшість представлених функцій, що використовуються для опису техніко-економічних показників, шляхом функціональних перетворень y по x (роздільно або одночасно) можуть бути зведені до лінійного вигляду. При цьому метод перетворень залежить від форми зв'язку.

Гіпербола вигляду $\bar{y} = \frac{k}{x} + b$ перетвориться в лінійну шляхом заміни.

Статична функція вигляду $\bar{y} = b k^x$ перетвориться в лінійну шляхом логарифмуванням.

У результаті маємо $\lg \bar{y} = \lg b + k \lg x$. Позначимо $\bar{y}' + \lg \bar{y}, b' = \lg b, x' = \lg x$.

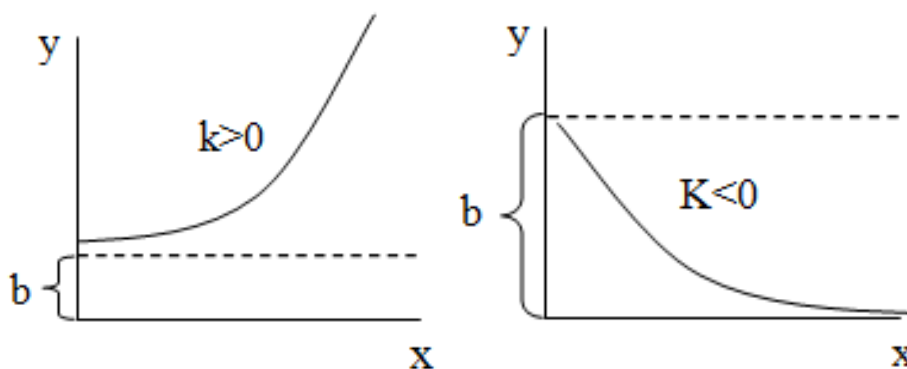


У результаті маємо $\bar{y}' = kx' + b'$.

Показова функція виду $y = be^{kx}$ перетвориться в лінійну логарифмуванням $\lg \bar{y} = \lg b + kx \lg e$.

Позначимо $\bar{y}' = \lg y'$, $b' = \lg b$, $k' = k \lg e$, при цьому $\lg e = 0.4343$.

У результаті маємо $y_1 x = b_1$.



Теоретична лінія регресії може бути подана у вигляді плавної кривої, яка кількісно виражає зв'язок між середніми інтервальними значеннями \bar{y}_{cp} і відповідними значеннями x (аргументами). Процес знаходження невідомих параметрів теоретичної залежності є однією з важливих проблем теорії кореляції і регресії.

Наприклад, при знаходженні параметрів параболи виду

$$\bar{y} = ax^2 + bx + c$$

необхідно скласти і вирішувати систему з трьох нормальних рівнянь, яка розв'язується, виходячи з вимоги методу найменших квадратів, тобто:

$$\varepsilon_i = \sum_{i=1}^n (y_i - \tilde{y}_i)^2 = \min.$$

Підставляючи

$$\tilde{y} = ax_i^2 + bx + c,$$

маємо

$$\varepsilon_i = \sum_{i=1}^n [y_i - (ax_i^2 + bx_i + c)]^2 = \min = F(a, b, c). \quad (4.61)$$

Знаходимо часткові похідні

$$\frac{dF}{da}, \frac{dF}{db}, \frac{dF}{dc}$$

і прирівнюємо їх до нуля

$$\begin{aligned} \frac{dF}{da} &= -2 \sum_{i=1}^n [y_i - (ax_i^2 + bx_i + c)] x_i^2 = 0, \\ \frac{dF}{db} &= -2 \sum_{i=1}^n [y_i - (ax_i^2 + bx_i + c)] x_i = 0, \\ \frac{dF}{dc} &= -2 \sum_{i=1}^n [y_i - (ax_i^2 + bx_i + c)] = 0. \end{aligned} \quad (4.62)$$

Після відповідного перетворення маємо

$$\begin{aligned} \sum_{i=1}^n y_i x_i^2 &= a \sum_{i=1}^n x_i^4 + b \sum_{i=1}^n x_i^3 + c \sum_{i=1}^n x_i^2, \\ \sum_{i=1}^n y_i x_i &= a \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i, \\ \sum_{i=1}^n y_i &= a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i + c_n. \end{aligned} \quad (4.63)$$

В узагальнених економетричних моделях визначаються також коефіцієнти кореляції, детермінації, критерії адекватності (t-статистики, F-критерій Фішера, перевіряється на мультиколінеарність, гетероскедастичність, автокореляцію залишків).

4.4 Економетричні моделі динаміки в оцінці нерухомого майна

Динамічні процеси, які здійснюються у системах, проявляються у вигляді ряду послідовно розташованих в хронологічному порядку значень того чи іншого показника, який в своїх вимірах відображає хід розвитку відповідного явища.

У аспекті дослідження динамічних процесів і побудови економетричних моделей динаміки розглядають такі поняття.

Динамічним рядом або рядом динаміки є послідовність спостережень одного показника, упорядкованих в залежності від послідовно зростаючих або спадаючих значень другого показника [69].

Часовий ряд – це послідовність спостережень $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$, кожне з яких відноситься до деякого відрізка часу t_1, t_2, \dots, t_n , або визначає результати за деякий період часу [70].

Складовими елементами рядів динаміки є цифрові значення показника, який називається *рівнем рядів* [70].

Часові ряди створені показниками, які характеризують відповідні явища, на визначені моменти часу, називаються *моментними* [70].

Приклад цього ряду представлено в таблиці 4.7.

Таблиця 4.7 – Чисельність робітників підприємства

Дата	1.01	1.02	1.03	1.04	30.04
Чисельність робітників	5600	5900	5400	5700	6000

Якщо рівні часового ряду створюються шляхом агрегування часового ряду за визначений проміжок часу, то такі ряди називаються інтервальними часовими рядами (табл. 4.8).

Таблиця 4.8 – Фонд заробітної плати робітників підприємства

Місяць	Січень	Лютий	Березень	Квітень
Фонд заробітної плати робітників, тис. грн.	88978,2	94521,1	96219,3	95310,9

Часові ряди можуть бути створені як із абсолютних значень показників, так і із середніх або відносних величин – це *похідні ряди* (табл. 4.9).

Таблиця 4.9 – Середньомісячна заробітна плата робітників підприємства

Місяць	Січень	Лютий	Березень	Квітень
Середня заробітна плата робітників, грн.	2300	2350	2410	2562

Під довжиною часового ряду розуміють час, який пройшов від початкового моменту спостереження до кінцевого. В наведених таблицях довжина всіх рядів дорівнює чотирьом місяцям.

Тренд – це рівняння $Y = d(t)$, що виражає в середньому зміну в часі показника, заданого рядом динаміки. Таке рівняння можна розглядати як апроксимацію часового ряду або як окремий випадок регресії. У зв'язку з цим математична динамічна модель, в якій розвиток модельованої економічної системи відображається у вигляді тренду її основних показників, називається *трендовою моделлю* [68].

У часових рядах економічних процесів можуть мати місце більші або менші регулярні коливання. Якщо вони носять строго періодичний або близький до нього характер і закінчуються протягом одного року, то їх називають *сезонними коливаннями*. В таких випадках, коли період коливань складає декілька років, то спостерігається *циклічна складова*.

Тренд, сезонна і циклічна складові називаються *регулярними, або систематичними складовими часового ряду* [70].

Складова частина часового ряду, яка залишається після виділення із нього регуляторних компонент, називається *випадковою, нерегулярною компонентою* [70].

Якщо систематичні компоненти часового ряду визначені вірно, то *залишкова послідовність* після виділення із часового ряду цих компонент *буде компонентою ряду*. Ця компонента має такі властивості:

- випадковість коливань рівня залишкової послідовності;
- відповідність розподілу випадкової компоненти нормальному закону розподілу;
- рівність математичного очікування випадкової компоненти нулю;
- незалежність значень рівнів випадкової послідовності, тобто відсутністю автокореляції.

Перевірка адекватності трендових моделей базується на перевірці виконання в залишковій послідовності вказаних чотирьох властивостей. Якщо не виконується одне з них, то модель визнається неадекватною; при виконанні всіх чотирьох властивостей модель адекватна.

Аналіз часових рядів є важливим етапом оцінки динаміки економічних процесів і побудові економетричних моделей динаміки.

Прикладами часових рядів також є щомісячна, щоквартальна, щорічна собівартість перевезення пасажирів, обсяг пасажирів, що перевозяться по депо, або рівнянню в цілому. Вихідні дані слід формувати по кожному з об'єктів у зв'язку з тим, що інформація буде більш достовірною ніж по групі об'єктів.

Маючи в розпорядженні свій часовий ряд для досліджуваного показника і для всіх чинників, необхідно перш за все виявити загальну тенденцію зміни цих величин (тренд, еволюційну складову, лінію рівняння).

Як показує дослідження економічних часових рядів, в них завжди міститься загальна тенденція, яку необхідно виявити. Співвідношення $Y = d(t)$ можна відшукати безпосередньо за звітними або дослідженими даними.

Використання ковзаючих середніх доцільно в разі достатньо довгого ряду. Число членів ковзаючої середньої повинно бути обумовлено міркуваннями за суттю процесу і залежно від кроку часового ряду.

Під час згладжування за допомогою ковзаючих середніх доводиться втрачати частину даних: при тричленному вирівнюванні – дві сторони таблиці, при чотирьох і п'яти членному вирівнюванні – відповідно три і чотири рядки. Якщо число даних не вірне, то таке скорочення даних навряд чи буде доцільним.

Питання про доцільну довжину часового ряду досить складне. З одного боку, як і завжди при пошуку апроксимуючої формули або рівняння регресії, виникає природне прагнення до збільшення масиву спостережень з метою підвищення точності надійності результатів, з другого боку, при обробці часових рядів слід врахувати небажаність використання старих даних. Приймати ці суперечливі вимоги можна тільки за рахунок зменшення довжини інтервалів часового ряду – скорочуючи крок ряду (шляхом переходу, наприклад від квартальних даних до місячних, від місячних до тижневих, і т.п. якщо такі дані за матеріалами звітності можна мати).

Розглянемо приклад. Дані собівартості пасажироперевезень міським електричним транспортом, представлені в таблиці 4.10, вирівняти за ковзаючою середньою і побудувати графік.

Розв'язання

Таблиця 4.10 – Статистичні дані собівартості пасажироперевезень по депо

t	Собівартість С, коп.	Тричленні суми	Тричленні ковзані середні	Чотиричленні суми	Чотиричленні ковзані середні	П'ятичленні суми	П'ятичленні ковзані середні
1	65,9	-	-	-	-	-	-
2	66,9	201,3	67,3	269,4	67,3	-	-
3	69,1	203,5	67,8	271,7	67,9	338,2	67,6
4	67,5	204,8	68,2	274,7	68,6	339,3	6,8
5	68,2	205,6	68,5	276,5	69,1	345,6	69,1
6	69,9	209,0	69,5	281,8	70,4	349,3	69,7
7	70,9	213,6	71,2	285,7	71,4	358,5	71,7
8	72,8	215,8	71,9	288,5	72,1	361,4	72,2
9	72,1	217,7	72,5	290,5	72,6	363,7	72,7
10	72,8	217,7	72,5	290,9	72,7	365,7	72,9
11	72,8	278,8	72,9	-	-	-	-
12	73,1	-	-	-	-	-	-

Як бачимо, середні дані більш наочно виражають основну тенденцію собівартості перевезення пасажирів.

Вихідні дані, що ковзають, подані на рисунку 4.1.

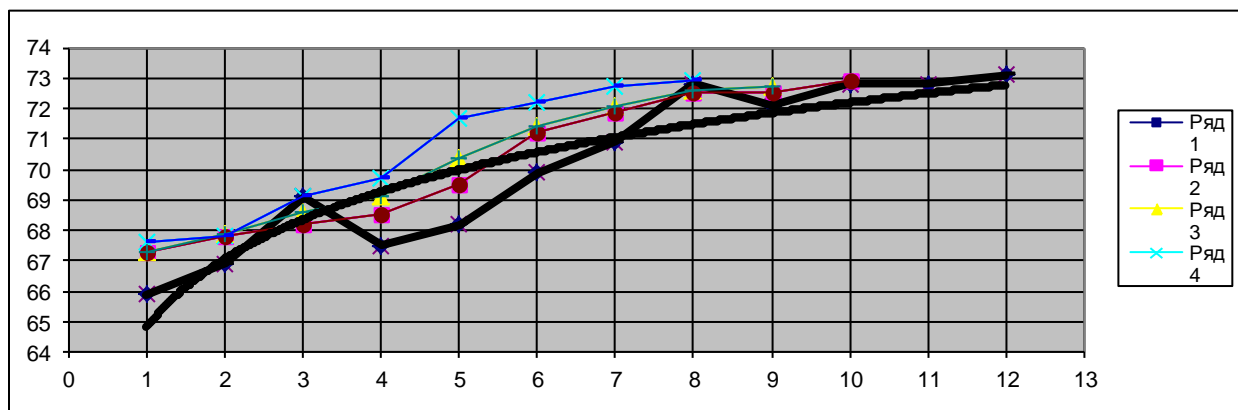


Рисунок 4.1 – Вихідні дані, ковзаючі і вирівнююча парабола:

1 – вихідні дані;

4 – п'ятичленні середні;

2 – тричленні ковзані;

5 – вирівнююча парабола.

3 – чотиричленні ковзані;

При визначенні загальної тенденції виникає два завдання:

- вибір форми рівняння, тобто вид функції $d(t)$;
- обчислення параметрів рівняння.

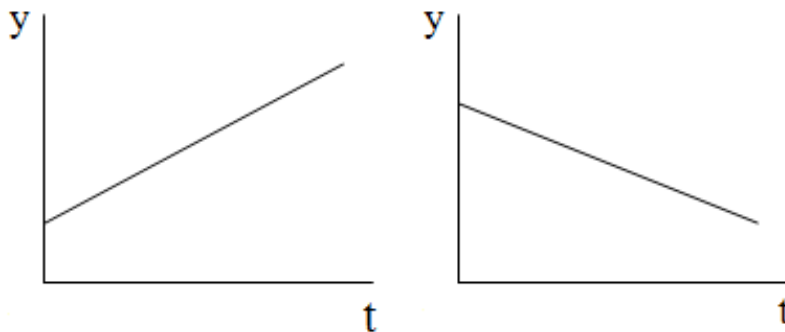
Слід зазначити, що аналіз часових рядів спрямований не тільки на визначення загальної тенденції і побудови моделі динаміки, а й на прогнозування економічних показників.

При виборі форми рівняння необхідно, як в статистичному регресійному аналізі, добре знати процес по суті. Так, для короткострокового прогнозування механіко-економічних показників найкращою формою тренда є опис зростання за законом складних процесів, для більш тривалого періоду прогнозування по цілому ряду показників – експонента з насиченням. Якщо ж сутність процесу не вимагає певної форми управління, то вибір проводиться за якнайменшою залишковою дисперсією.

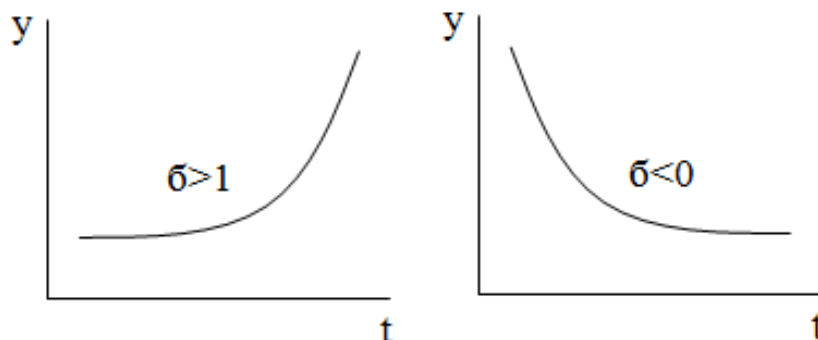
Графічне подання часового ряду також допоможе в цьому виборі.

Практика показує, що доцільно піддавати випробуванню залишкову дисперсію по чотирьох монотонних функціях:

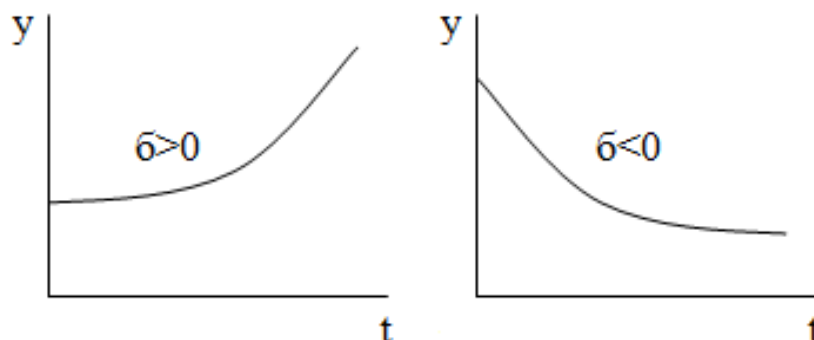
1. Лінійній $\bar{Y}_t = at + b$.



2. Степеневій $\bar{Y} at^\alpha$ або $\ln y_t + \alpha \ln t$.

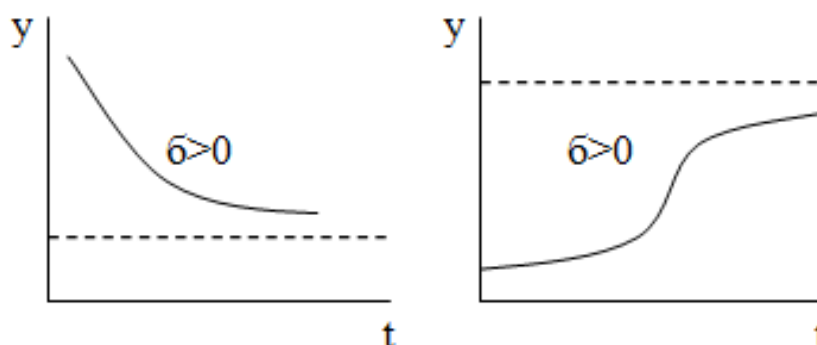


3. Експоненціальний $\bar{Y} at^\alpha$ або $\ln y_t = \ln a + \alpha t$.



4. Експоненти з насиченням $\bar{Y} ae^{\frac{\alpha}{t}}$ або $\ln y_t = \ln a + \frac{\alpha}{t}$.

При цьому відхилення від тренду визначаються відповідно у вигляді:



$$E_t = Y_t - (a + kt),$$

$$E_t = \ln y_t - (\ln a + \alpha \ln t),$$

$$E_t = \ln y_t - (\ln a + \alpha t), \quad (4.64)$$

$$E_t = \ln y_t - (\ln a + \frac{\alpha}{t}).$$

Всі параметри α знаходять за методом якнайменших квадратів, що приводить до системи нормальних рівнянь:

$$\begin{cases} \sum_{t=1}^n Y_t = \alpha \sum_{t=1}^n t + na, \\ \sum_{t=1}^n tY_t = \alpha \sum_{t=1}^n t^2 + a \sum_{t=1}^n t; \end{cases} \quad (4.65)$$

$$\begin{cases} \sum_{t=1}^n \ln Y_t = \alpha \sum_{t=1}^n \ln t + n \ln a, \\ \sum_{t=1}^n \ln t \ln y_t = \alpha \sum_{t=1}^n \ln t^2 + \ln a \sum_{t=1}^n \ln t; \end{cases} \quad (4.66)$$

$$\begin{cases} \sum_{t=1}^n \ln Y_t = \alpha \sum_{t=1}^n t + n \ln a, \\ \sum_{t=1}^n t \ln y_t = \alpha \sum_{t=1}^n t^2 + \ln a \sum_{t=1}^n t; \end{cases} \quad (4.67)$$

$$\begin{cases} \sum_{t=1}^n \ln Y_t = \alpha \sum_{t=1}^n \frac{1}{t} + n \ln a, \\ \sum_{t=1}^n \frac{1}{t} \ln y_t = \alpha \sum_{t=1}^n \frac{1}{t^2} + \ln a \sum_{t=1}^n \frac{1}{t}. \end{cases} \quad (4.68)$$

Знайшовши $\min \delta_y^2$ для відповідної залежності, знаходять функцію, яка в порівнянні з іншими найкраще апроксимує початковий часовий ряд.

Використання багатопараметричних функцій з метою апроксимації недоцільне. Хоча за допомогою таких функцій можна отримати добре наближення вихідним даним, але, таким чином математично описується не стільки загальна тенденція, скільки випадкові від неї відхилення; з'являються не виправдані особливості процесу – максимуми і мінімуми.

Крім того, складання таких функцій і їх застосування для практичних розрахунків різко ускладнюється.

Розглянемо приклад. Для вищеперерахованих даних, використовуючи степеневу залежність $Y_t = at^\alpha$, розрахуємо її параметри.

Розв'язання

Для визначення параметрів рівняння представимо в таблиці 4.11 необхідні розрахунки.

Таблиця 4.11 – Розрахунок статистичних характеристик рівняння

	Y_t	$\ln t$	$\ln t^2$	$\ln Y_t$	$\ln Y_t \ln t$	$\ln Y_t$	Y_t	ε	ε^2
1	2	3	4	5	6	7	8	9	10
1	65,9	0,00	0,00	1,8189	0,00	1,8035	63,6	2,3	5,29
2	66,9	0,3010	0,0906	1,8254	0,5494	1,8211	66,24	0,66	0,44
3	69,1	0,4771	0,2276	1,8395	0,8774	1,8314	67,32	1,28	1,64
4	67,5	0,6021	0,3625	1,8293	1,1014	1,8387	68,98	-1,48	2,19
5	68,2	0,6990	0,4886	1,8331	1,2813	1,8443	69,88	-1,68	2,82

Продовження таблиці 4.11

1	2	3	4	5	6	7	8	9	10
6	69,9	0,7782	0,6056	1,8445	1,4354	1,8490	70,13	0,73	0,53
7	70,9	0,8451	0,7142	1,8505	1,5639	1,8524	71,19	-0,29	0,08
8	72,9	0,9031	0,8156	1,8627	1,6822	1,8566	71,88	1,02	1,04
9	72,1	0,9542	0,9109	1,8579	1,7728	1,8593	72,33	-0,23	0,05
10	72,8	1,000	1,00	1,8621	1,8626	1,8620	72,88	-0,08	0,00
11	72,8	1,0414	1,0845	1,8621	1,9392	1,8644	73,16	-0,36	0,13
12	73,2	1,0792	1,1647	1,8645	2,0122	1,8666	73,52	-0,32	0,12
Σ		8,6824	7,4648	22,1506	16,0959				14,32

Система нормальних рівнянь має вигляд:

$$\begin{cases} \sum_{i=1}^n \ln Y_t = \alpha \sum_{i=1}^n \ln t + n \ln a, \\ \sum_{i=1}^n \ln t \ln Y_t = \alpha \sum_{i=1}^n \ln t^2 + \ln a \sum_{i=1}^n \ln t. \end{cases} \quad (4.69)$$

Підставивши відповідні значення з таблиці 4.11, отримаємо:

$$\begin{cases} 22,1506 = 8,6824\alpha + 12 \ln a, \\ 16,0959 = 7,4648\alpha + 8,6424 \ln a. \end{cases} \quad (4.70)$$

Вирішивши систему рівнянь, одержимо:

$$\alpha = 0,0585; \quad \ln a = 1,8035.$$

Маємо рівняння:

$$\overline{\ln Y_t} = \ln a + \alpha \ln t \rightarrow \overline{\ln Y_t} = 1,8035 + 0,0585 \ln t$$

або

$$Y_t - at^\alpha \rightarrow \overline{Y_t} = 63,60t^{0,0585} \quad (4.71)$$

Прогноз на 13 і 14 періоди складе:

$$Y_{13}=72,83; \quad Y_m=73,02.$$

Середній квадрат відхилення вихідних значень від розрахункових (дисперсія) матиме вигляд:

$$\delta_y^2 = \frac{1}{n} \sum_{t=1}^n \varepsilon^2 = \frac{14,33}{12} = 1,194, \quad (4.72)$$

а середнє квадратичне відхилення

$$\delta_y \sqrt{1,194} = 1,092,$$

що в порівнянні з середнім розміром складає

$$\tilde{y} \frac{1,092}{71,11} * 100\% = 1,53\% .$$

Проте зазначимо, похибки апроксимації особливо великі на кінцях базисного періоду, що обумовлюють велику помилку прогнозу. Можна сказати, що залежність підібрана невдало.

Якщо протягом базисного періоду процес, що вивчається, суттєво змінився в результаті появи нових чинників (сезонні коливання), то для апроксимації часового ряду слід скористатися двома або більш окремими аналітичними виразами, розглядаючи їх як частини науково-безперервної функції. При цьому прогнозування проводиться за останньою дугою і необхідно уточнити, який допустимий інтервал прогнозування. Факт істотності змін для показника слід встановлювати як якісно, так і статистично.

Можна скористатися і графічним способом: побудувавши три тренди по кожному періоду в цілому по всьому ряду.

Статистична перевірка може бути здійснений таким прикладом дисперсійного аналізу. Нехай значення показника до і після деякого моменту задані рядами:

$$Y_{11}, Y_{12}, \dots, Y_{1h}, \quad (4.73)$$

$$Y_{21}, Y_{22}, \dots, Y_{2h} \quad (4.74)$$

з середніми значеннями і дисперсіями, визначеними по формулах:

$$\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij},$$

$$\sigma^2 = \frac{1}{n_i} \sum_{j=1}^{h_i} (y_{ij} - \bar{y}_j)^2, \quad j = 1, 2. \quad (4.75)$$

Обчислюємо загальну середню і загальну дисперсію з'єднаного ряду:

$$\bar{y} = \frac{1}{h} \sum_{i=1,2}^{h_i} \sum_{j=1}^{h_i} y_{ij},$$

$$\sigma_y^2 = \frac{1}{h-1} \sum_{i=1,2}^{h_i} (y_{ij} - \bar{y})^2. \quad (4.76)$$

Розчленувавши повну дисперсію ряду на частини, одержуємо:

$$\sum_{i=1,2}^{h_i} \sum_{j=1}^{h_i} (y_{ij} - \bar{y})^2 = \sum_{i=1,2}^{h_i} \sum_{j=1}^{h_i} (y_{ij} - \bar{y}_i)^2 + \sum_{i=1,2}^{h_i} (\bar{y}_i - \bar{y})^2 h_i. \quad (4.77)$$

Враховуючи число ступенів свободи кожної з сум в рівнянні, одержуємо:

$$S_i^2 = \frac{\sum_{i=1,2}^{h_i} \sum_{j=1}^{h_i} (y_{ij} - \bar{y}_i)^2}{n-2},$$

$$S_2^2 = \frac{\sum_{i=1,2}^{h_i} (\bar{y}_i - \bar{y})^2 h_i}{n-2}. \quad (4.78)$$

Відношення $\frac{S_2^2}{S_1^2}$ порівнюємо з відповідним значенням розподілу Фішера.

Якщо $\frac{S_2^2}{S_1^2} < F_{5\%} [1, n-2]$ при рівні значущості 5 % вважаємо періоди, що

вивчаються, не істотно різними у значенні даного показника Y .

Якщо $\frac{S_2^2}{S_1^2} < F_{5\%} [1, n-2]$ при рівні значущості 1 % вважаємо періоди, що

вивчаються, суттєво різними за показником Y і будуємо тренд з двох частин, різних тільки за параметрами або видом функції $d(t)$.

Розглянемо приклад. Методику обробки рядів динаміки за наявності сезонних коливань можна проілюструвати на прикладі собівартості пасажироперевезень за період 2002 – 2007 роки.

Розв'язання

Виявлення загальної тенденції на підставі даних таблиці 4.12 починаємо з побудови графіка.

Таблиця 4.12 – Динаміка статистичних показників

Роки	t	Значення показника, С, коп.	Квартал			
			I	II	III	IV
2002	1	58,71	62,3	56,88	59,34	56,72
2003	2	60,13	62,78	58,35	60,84	58,78
2004	3	60,83	63,47	57,88	62,58	59,4
2005	4	65,70	69,52	63,02	63,89	66,51
2006	5	66,08	67,23	62,99	65,65	67,54
2007	6	66,76	68,59	60,56	66,00	68,29

У цьому прикладі (рис. 4.2) спостерігається різкий перелом характеру змін в 2004 році. Тому неможливо підібрати єдину математичну функцію зростання, задовільно апроксимуючу дані про собівартість пасажироперевезень за всі роки. У зв'язку з цим розбиваємо тимчасовий діапазон на дві частини – 2002 – 2005 роки і 2005 – 2007 роки.

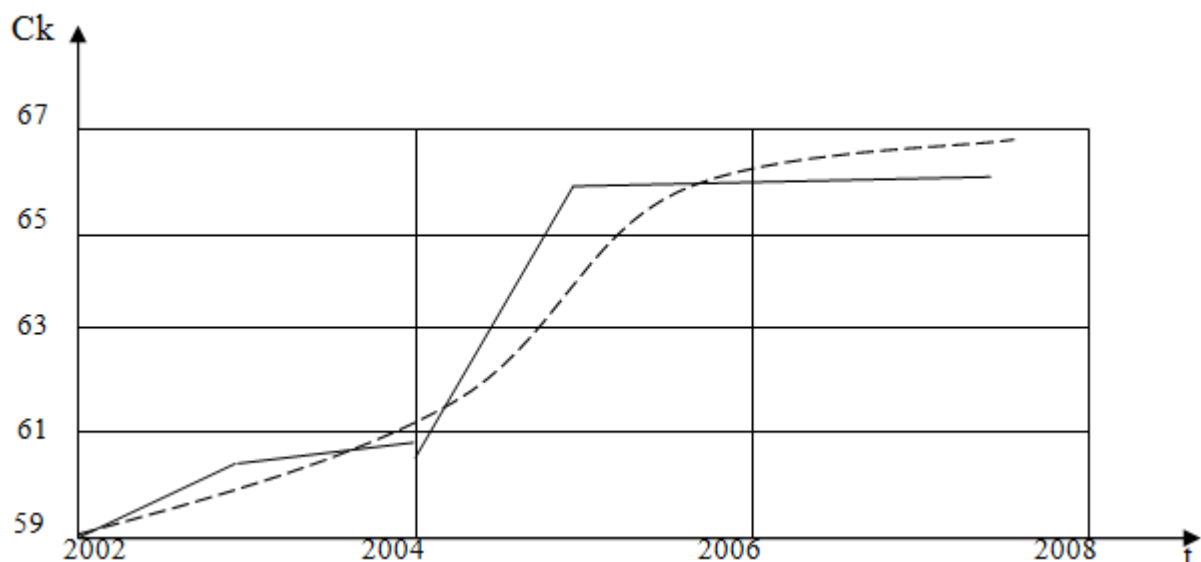


Рисунок 4.2 – Динаміка собівартості пасажироперевезень

Для першого ряду підбираємо експоненту, для другого – експоненту з насиченням. При визначенні параметрів рівнянь використовуємо розрахунки, що зведені відповідно в таблиці 4.13 і 4.14.

Таблиця 4.13 – Розрахунок параметрів експоненти

Роки	Y	ℓ_{ny}	t	$V = \ell_{ny-1,78}$	Vt	t^2	\bar{V}	$\frac{\ell_{ny=1,78+}}{\bar{V}}$	\bar{Y}	ε	ε^2	$\beta\%$
2002	58,71	1,7761	0	-0,0039	0	0	-0,013	1,7787	60,08	-1,37	1,88	2,28
2003	60,13	1,7791	1	-0,0004	-0,004	1	0,0021	1,7821	60,53	-0,4	0,16	0,66
2004	60,83	1,7841	2	0,0041	0,0082	4	0,0155	1,7955	62,44	-1,61	2,59	2,57
2005	65,70	1,8176	3	0,0376	0,1182	9	0,0289	1,8089	64,40	1,3	1,69	2,01
Σ			6	0,0374	0,1206	14					6,32	

Таблиця 4.14 – Розрахунок параметрів експоненти з насиченням

Роки	Y	ℓ_{ny}	t	$t^1=t-2$	$V = \ell_{ny-1,82}$	$\frac{1}{t^1}$	$\frac{V}{t^1}$	$\frac{1^2}{t^1}$	\bar{V}	$\ell_{ny=1,82+V}$	\bar{Y}	ε	ε^2	$\beta\%$
2005	65,7	1,8176	3	1	0,0024	1	-0,024	1	-0,0084	1,8116	64,8	0,9	0,81	1,388
2006	65,08	1,82	4	2	0,000	0,50	0,00	0,250	0,0033	1,8233	66,58	-0,5	0,25	0,75
2007	66,76	1,8245	5	3	0,0045	0,333	0,0015	0,111	0,0072	1,8272	67,16	-0,4	0,16	0,59
Σ					0,0021	1,833	-0,0009	1,361					1,22	

Системи нормальних рівнянь:

$$\begin{cases} 0,0374 = 6\alpha + 4 \ln a \\ 0,1206 = 14\alpha + 6 \ln a \end{cases},$$

$$\begin{cases} 0,0021 = 1,833\alpha^1 + 3 \ln a^1 \\ -0,0009 = 1,361\alpha^1 + 1,833 \ln a^1 \end{cases}. \quad (4.79)$$

Звідки

$$\ln a = -0,0113,$$

$$a = 0,0134,$$

$$\ln_a^1 = 0,015,$$

$$a^1 = -0,0234,$$

$$V = -0,0113 + 0,0134 \ t,$$

$$V^1 = 0,015 - \frac{0,0234}{t}.$$

Тоді одержуємо:

$$\ln \bar{Y}_t = 1,7687 + 0,0134 \ t,$$

$$\ln \bar{Y}_t = 1,835 - \frac{0,0234}{t},$$

$$\bar{Y} = 58,7^{0,0134t}, \ 0 \leq t \leq 3,$$

$$\bar{Y}_t = 68,39 - \frac{0,034}{t}, \ 3 \leq t \leq 5,$$

$$\max \beta\% = 2,58,$$

$$\max \beta\% = 1,388.$$

Апроксимація цілком задовільна.

Для 2005 року приймаємо значення собівартості:

$$\bar{Y} = \frac{64,40 + 64,80}{2} = 64,60.$$

Прогноз на 2004 рік. При $t=6$:

$$\ln y = 1,835 - \frac{0,0234}{4} = 1,829,$$

$$Y_{2004} = 67,49.$$

Проте, через сезонні коливання прогнозування за сумарними річними даними є абсолютно недостатнім. Тому необхідно прогнозувати за окремими періодами, у даному прикладі за вихідними квартальними даними.

При оцінці динаміки процесів і їх прогнозуванні необхідно спиратись на обґрунтовану теорію, що встановлює правомочність оцінки і прогнозування за допомогою моделі і помилки вірогідності прогнозу. Оцінка такої помилки за допомогою функції зростання неможлива, тому особливий інтерес представляють авторегресійні моделі.

Авторегресією називається рівняння, що визначає змінну x в момент t (або t -й період) через її значення в попередні періоди: $(t-1) \cdot (t-2) \cdot \dots \cdot (t-k)$. Лінійне авторегресійне рівняння записуємо у вигляді:

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_k x_{t-k}. \quad (4.80)$$

Першим етапом дослідження часового ряду змінної x є виділення загальної тенденції у вигляді функції $d(t)$ і визначення залишків ε_t у формі $\varepsilon_t = x_t - d(t)$ чи $\varepsilon_t = d(x_t)$.

Якщо залишки ε_t незалежні, тобто не можуть бути представлені як функція часу, то функція $d(t)$ охоплює повністю еволюційну складову змінної x_t . При цьому залишається знайти закон їх розподілу ε_t і, прийнявши гіпотезу про збереження цього закону розподілу на прогнозований період, побудувати довірчий інтервал для прогнозованої величини x_t за функцією $d(t)$.

Якщо ж залишки ε_t залежні, тобто містять деяку тенденцію, то її можна виявити за допомогою коефіцієнта автокореляції.

Проводячи зсув значень ε_t на один рядок, останнє значення переміщаємо на перше місце, одержуємо таблицю 4.15.

Таблиця 4.15 – Залишки змінних ряду динаміки

ε_t	ε_{t-1}
ε_1	ε_n
ε_2	ε_1
ε_3	ε_2
.....
ε_n	ε_{n-1}

Обчислюємо циклічний коефіцієнт кореляції між рядами ε_t і ε_{t-1} за формулою:

$$r(\varepsilon_t, \varepsilon_{t-1}) = \frac{\sum \varepsilon_t * \varepsilon_{t-1}}{\sum \varepsilon_t^2}. \quad (4.81)$$

Формула (4.81) виходить із звичайної формули для визначення коефіцієнта кореляції, якщо покласти:

$$\sum \varepsilon_t = \sum \varepsilon_{t-1} = 0, \quad (4.82)$$

$$\sum (\varepsilon_{t-1})^2 = \sum (\varepsilon_t)^2. \quad (4.83)$$

Формула (4.82) виходить з того, що параметри функції $d(t)$ визначаються за методом найменших квадратів, а формула (4.83) – з циклічної таблиці 4.15.

Аналогічно, зсовуючи ε_t на 2, 3, ..., k рядків, одержуємо циклічну таблицю послідовних відхилень.

Таблиця 4.16 – Циклічна таблиця послідовних відхилень

t	ε_t	ε_{t-1}	ε_{t-2}	ε_{t-k+1}	ε_{t-k}
1	ε_1	ε_n	ε_{n-1}		ε_{n-k+2}	ε_{n-k+1}
2	ε_2	ε_1	ε_n		ε_{n-k+3}	ε_{n-k+2}
3	ε_3	ε_2	ε_1		ε_{n-k+4}	ε_{n-k+3}
....
k	ε_k	ε_{k-1}	ε_{k-2}		ε_1	ε_n
k+1	ε_{k+1}	ε_k	ε_{k-1}		ε_2	ε_1
k+2	ε_{k+2}	ε_{k+1}	ε_k		ε_3	ε_2
.....
n	ε_n	ε_{n-1}	ε_{n-2}		ε_{n-k+1}	ε_{n-k}

За даними таблиці 4.16 визначаємо всі циклічні коефіцієнти автокореляції:

$$r(\varepsilon_{x_t} \varepsilon_{x_{t-j}}) = \frac{\sum_t \varepsilon_{x_t} \varepsilon_{x_{t-j}}}{\sum_t (\varepsilon_t)^2}, \quad j = 1, 2, \dots, k, \quad (4.84)$$

$$r(\varepsilon_{x_{t-1}} \varepsilon_{x_{t-j}}) = \frac{\sum_t \varepsilon_{x_{t-1}} \varepsilon_{x_{t-j}}}{\sum_t (\varepsilon_{x_t})^2}. \quad (4.85)$$

Циклічний коефіцієнт автокореляції не підпорядковується нормальному закону розподілу, його розподіл асиметричний, суттєві величини коефіцієнтів автокореляції при певному рівні значущості різні для додатних і від'ємних його значень. Рівні значущості коефіцієнтів автокореляції 5 % і 1 % подані в спеціальних таблицях.

Знайдені значення

$$r_1, r_2, \dots, r_{n-k-1}$$

перевіряємо по таблиці 5 % і 1 % рівнів вірогідності коефіцієнтів автокореляції.

Якщо

$$|r_{\text{сам}}(n)| < |r_{5\%}(n)|,$$

то приймаємо гіпотезу про неавтокорельованість залишків ε_t , якщо

$$|r_{\text{сам}}(n)| > |r_{1\%}(n)|,$$

то відкидаємо гіпотезу про їх неавтокорельованість.

За циклічними коефіцієнтами автокореляції складаємо матрицю і її обертаємо. Як і в разі звичайної регресії, перевіряємо наявність мультиколінеарності кожного з чинників $\varepsilon_{x_{t-j}}$, $j = 1, 2, \dots, k$ від сукупності інших і зберігаємо тільки лінійно незалежні аргументи.

Будуємо лінійну авторегресійну модель:

$$\varepsilon_t = a_1 \varepsilon_{t-1} + a_2 \varepsilon_{t-2} + \dots + a_k \varepsilon_{t-k}, \quad (4.86)$$

що виражає ε_t в період t за допомогою значень

$$\varepsilon_{t-j}, j = 1, 2, \dots, k$$

за k попередніх періодів.

При цьому в рівнянні повинні бути збережені тільки суттєві і лінійно незалежні коефіцієнти.

Якщо виявляються a_j -коефіцієнти, що не задовольняють вказаним вимогам, то модель потребує перерахунку (починаючи з розрахунку автокореляційної матриці більш низького порядку).

Оскільки параметри рівняння тренда визначали за методом найменших квадратів, то в разі його коректного підбору відповідні відхилення підкоряються нормальному розподілу, і рівняння регресії можна відшукати з лінійної форми:

$$\ln x_t = a_1 \ln x_{t-1} + a_2 \ln x_{t-2} + \dots + a_k \ln x_{t-k} + F(t), \quad (4.87)$$

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_k x_{t-k} + F(t). \quad (4.88)$$

Яким повинне бути число членів рівняння? Це питання слід вирішувати в поєднанні професійних вимог процесу, що по суті вивчається за математико-статистичними критеріями. Так, якщо статистичний ряд містить тижневі дані, то особливий інтерес до чотиричленної моделі залежності рівня показника від тижневих рівнів за весь попередній місяць. У разі місячних даних цікава тричленна авторегресія, а для даних, зібраних по роках, – п'ятичленна.

Статистичні критерії можуть встановити відсутність автокорельованості залишків із віднімання з табличних значень ε_t їх розрахункових значень:

$$\eta_t = \varepsilon_t - (a_1 \varepsilon_{t-1} + a_2 \varepsilon_{t-2} + \dots + a_k \varepsilon_{t-2k}). \quad (4.89)$$

Існує декілька статистичних критеріїв. Один з них заснований на порівнянні середнього квадрата послідовних різниць η_t :

$$\frac{1}{n-k-1} \cdot \frac{1}{n-k-1} \sum_{i=k+2}^n (\eta_t - \eta_{t-1})^2 \quad (4.90)$$

з дисперсією величини

$$\frac{1}{n-k} \cdot \frac{1}{n-k} \sum_{i=k-1}^n (\eta_t)^2. \quad (4.91)$$

Складаємо відношення середнього квадрата послідовних різниць до середнього квадрата самих величин:

$$K = \frac{\frac{1}{n-k-1} \sum_{i=k+2}^n (\eta_i - \eta_{i-1})^2}{\frac{1}{n-k} \sum_{i=k+1}^n (\eta_i)^2}. \quad (4.92)$$

Якщо K_{stat} потрапляє в допустиму область при рівні значущості 5 %, а саме:

$$K_{5\%}(n-k) < K_{stat}(n-k) < K_{5\%}^1(n-k),$$

то приймаємо гіпотезу неавтокорельованості залишків η_t , а, отже, і достатності числа членів K авторегресійної моделі.

Якщо ж

$$K_{stat}(n-k) < K_{\%}(n-k) \text{ або } K_{stat} > K_{1\%}(n-k),$$

то відкидаємо гіпотезу неавтокорельованості залишків η_t і рахуємо число членів рівняння недостатності. У цьому випадку число членів рівняння треба збільшити, якщо довжина ряду дозволяє це.

Користуючись для прогнозу розробленими рівняннями, можна знайти довірчий інтервал для значення прогнозованого показника.

Якщо прогнозований показник рівний \tilde{x}_t , то розмір показника x_t запишемо у вигляді:

$$\tilde{x}_t - t \frac{1}{n} \sum_{t=1}^n \varepsilon_{x_t}^2 \leq x_t \leq +t \frac{1}{n} \sum_{t=1}^n \varepsilon_{x_t}^2. \quad (4.93)$$

Викладена методика складання авторегресійних моделей, використані критерії і побудований довірчий інтервал можна застосовувати тільки для великих вибірок, коли довжина ряду n не менше 30.

Помилка прогнозу з отриманих рівнянь визначається за дисперсією ε_t .

Оскільки

$$\tilde{x}_t - x_t = \varepsilon_t, \quad (4.94)$$

то

$$B_{сер} \{ |\bar{x}_t - x_t| = |\varepsilon_t| \leq t_\alpha \sigma_\varepsilon \} = P_\alpha, \quad (4.95)$$

де P_α – задана вірогідність, $P_\alpha = 1 - \alpha$;

t_α – відповідна межа по $C(n-k)$ ступеням свободи Ст'юдента:

$$\sigma_\varepsilon = \sqrt{\frac{\sum_{t=k+1}^n \varepsilon_t^2}{n-k}}. \quad (4.96)$$

Розглянемо *приклади* складання авторегресійних моделей.

Одночленна модель. Щомісячний пробіг рухомого складу міського електротранспорту на 1000 пасажирів, що перевозяться, заданий рядом в графі 2 таблиці 4.17. Знайдіть параметри одночленної авторегресійної моделі і спрогнозуйте щомісячний пробіг рухомого складу міського електротранспорту.

Розв'язання

Наявність експоненціального ряду (рисунок 4.3) дозволяє розраховувати на придатність одночленної моделі $\bar{x}_t = a_1 x_{t-1}$.

Система нормальних рівнянь для визначення параметра a_1 має вигляд:

$$\sum_{t=2}^{15} x_t x_{t-1} = a_1 \sum_{t=2}^{15} x_{t-1}^2. \quad (4.97)$$

З таблиці 4.17 (графи 4 і 5) виходить:

$$367673,4 = 364278,2 a_1.$$

Звідки

$$a_1 = \frac{367673,4}{364278,2} = 1,0087 \approx 1,01.$$

Одержуємо рівняння

$$\bar{x}_t = 1,01 \cdot x_{t-1}.$$

Обчислюємо значення

$$\bar{x}_t = 1,01 \cdot x_{t-1} \text{ (графа 6)}$$

і знаходимо значення

$$\varepsilon_t = x_t - x_{t-1} \text{ (графа 7)} \quad \sum \varepsilon_t = 9,4,$$

що несуттєві в порівнянні з розмірами x_t .

Обчислюємо коефіцієнт циклічної автокореляції r_1 . За графами 9 і 10 отримаємо:

$$r_1 = r(\varepsilon_t \cdot \varepsilon_{t-1}) = \frac{\sum_{t=k+1}^n \varepsilon_t \cdot \varepsilon_{t-1}}{\sum_{t=k+1}^n \varepsilon_t^2} = \frac{-184,14}{661,79} = -0,278. \quad (4.98)$$

З таблиці 4.17 знаходимо:

$$n^I = 15-1=14, r < 0, r_{5\%} = -0,479.$$

Оскільки $|r_1| < |r_{5\%}|$, то кореляція $\varepsilon_t, \varepsilon_{t-1}$ несуттєва.

Аналогічно за графами 12 і 10 (табл. 4.17) одержуємо:

$$r_2 = \frac{-275,68}{661,68} = 0,416,$$

що свідчить про несуттєвість кореляції $\varepsilon_t, \varepsilon_{t-2}$.

У даному випадку критерієм є критерій Дж. Неймана.

Обчислюємо різницю ε_t та ε_{t-1} за графою 13 і $(\varepsilon_t - \varepsilon_{t-1})^2$ за графою 14.

Одержуємо

$$K = \frac{\frac{1}{n-k-1} \sum_{t=k+2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\frac{1}{n-k} \sum_{t=k+1}^n \varepsilon_t^2} = \frac{1369,15/13}{661,79/14} = \frac{105,32}{47,27} = 2,23. \quad (4.99)$$

За таблицею 4.17 для $n_1 = 14$ рівень значущості $K_{5\%}$ рівний 1,2725 при $r > 0$ і 3,0352 у разі $r < 0$. Розрахунки свідчать, що коли в генеральній сукупності автокореляція між залишками ε_t відсутня, то в 95 % вибірок буде $K > 1,272$ у випадку $r > 0$ і $K < 3,0352$ при $r < 0$.

У даному прикладі значення K потрапляє в допустиму область при 5 % рівні значущості $K > 1,2725$. Отже, гіпотеза неавтокорельованості залишків ε_t стверджується і авторегресійне рівняння $x_t = 1,01 x_{t-1}$ приймається.

Помилка прогнозу при середньоквадратичному відхиленні:

$$\sigma_\varepsilon = \sqrt{\frac{\sum_{t=k+1}^n \varepsilon_t^2}{n-k}} = \sqrt{\frac{661,79}{14}} = 6,87. \quad (4.100)$$

Складаємо

$$B_{сер} \{ |x_t - \bar{x}_t| \leq t_\alpha * 6,87 \} = P_\alpha. \quad (4.101)$$

При 95 % гарантійній вірогідності $t_\alpha = 2,1$ за таблицею П.4 [19] і помилці прогнозу, що не перевищить 14,42 та складає приблизно 8 %:

$$\bar{x} - 14,42 \leq 1,01 x_{t-1} \leq \bar{x} + 14,42. \quad (4.102)$$

Визначаємо прогноз на 16-й і 17-й періоди з похибкою, що не перевершує 14,42 (рис. 4.3):

$$\bar{x}_{16} = 1,01 * x_{15} = 1,01 * 175,3 = 177,05, \quad (4.103)$$

$$\bar{x}_{17} = 1,01 * x_{16} = 1,01 * 177,05 = 178,78. \quad (4.104)$$

Таблиця 4.17 – Розрахунок параметрів одночленної авторегресійної моделі

t	x_t	x_{t-1}	x_t x_{t-1}	x_{t-1}²	$\tilde{x}_{t=1,01} x_{t-1}$	ε_t	ε_{t-1}	ε_t*ε_{t-1}	ε_t²	ε_{t-2}	ε_t*ε_{t-2}	ε_t-ε_{t-1}	(ε_t-ε_{t-1})²
1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	153,1	-	-	-	-	-	-	-	-	-	-	-	-
2	153,3	153,1	23470,2	23439,6	154,6	-1,3	5,4	-7,02	1,69	1,7	2,21	5,1	26,01
3	148,4	153,3	22749,7	23500,9	154,8	-6,4	-1,3	8,32	40,96	5,4	-34,56	-5,4	29,16
4	148,9	148,4	22096,8	22026,6	149,9	-1,0	-6,4	6,4	1,0	-1,3	1,3	-11,0	14,0
5	160,4	148,9	23883,6	22171,2	150,4	10	1,0	10,0	1,0	-6,4	-64,6	11,5	132,25
6	160,5	160,4	25744,2	25728,2	162,0	-1,5	10,0	-15,0	2,25	-1,0	1,5	33	10,89
7	157,3	160,5	25246,6	25760,2	162,1	-4,8	-1,5	7,2	23,04	10,0	-48,0	-16,5	272,25
8	170,6	157,3	26835,4	24743,3	158,9	11,7	-4,8	-16,5	136,89	-1,5	-17,55	3,3	10,89
9	163,9	170,6	27961,3	29104,4	172,3	-8,4	11,7	-98,27	70,56	-4,8	40,32	-7,4	54,76
10	164,3	163,9	26928,3	26863,2	165,3	-1,0	-8,4	8,4	1,0	11,7	-11,7	-16,1	259,21
11	170,9	164,3	28078,9	26994,5	155,8	15,1	1,0	15,1	228,01	-8,4	-126,84	19,8	292,04
12	167,9	170,9	28694,1	29206,8	172,6	-4,7	15,1	-70,97	22,09	-1,0	4,7	-3,2	10,24
13	168,1	167,9	28223,9	28190,4	169,6	-1,5	-4,7	7,05	2,25	15,1	-22,65	0,2	0,04
14	168,2	168,1	28274,4	28257,6	169,9	-1,7	-1,5	2,55	2,89	-4,7	7,99	-7,1	50,41
15	175,3	168,2	29485,5	28291,2	169,9	5,4	-1,7	9,18	29,16	-1,5	-8,4	-	-
16	175,5		367673,4	364278,2		9,4		-184,64	661,179		-275,68		1369,15

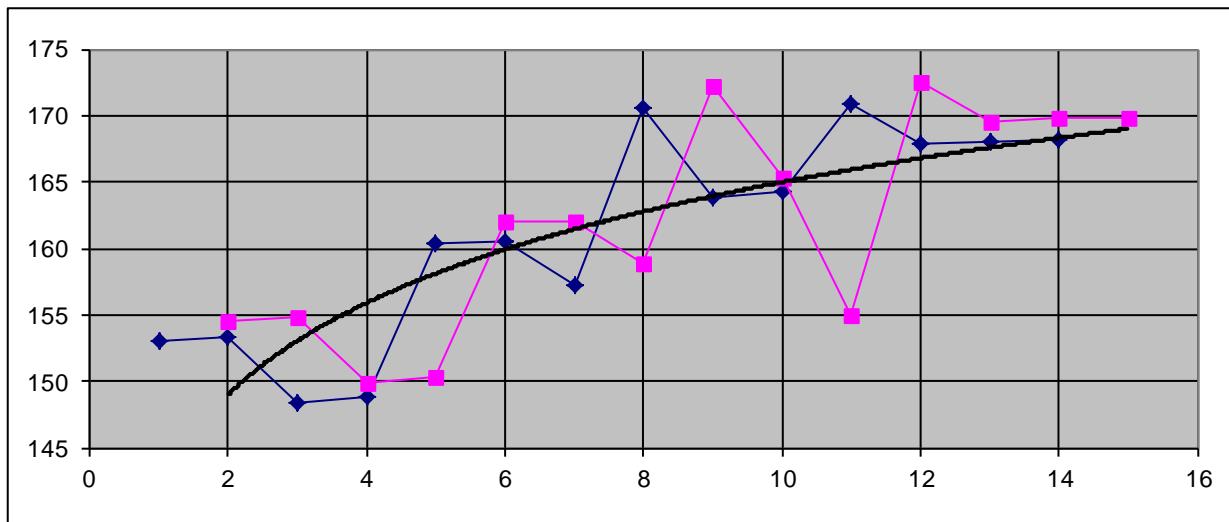


Рисунок 4.3 – Одночленна авторегресійна модель:

1 – вихідні дані; 2 – одночленна авторегресія;

3 – вирівнююча гіпербола

Багаточленна модель. Щомісячна реалізація цегли (в тисячах штук) на базі будівельних матеріалів за 18 місяців представлена в таблиці 4.18 (граф 2).

Потрібно скласти модель для прогнозування місячної потреби в цеглі на найближчі місяці.

Розв'язання

Розрахунок багаточленної авторегресійної моделі подано в таблиці 4.18.

Таблиця 4.18 – Розрахунок параметрів багаточленної авторегресійної моделі

t	x_t	x_{t-1}	$\bar{x}_t = 1,103x_{t-1}$	$\varepsilon_t = x_t - x_{t-1}$	ε_{t-1}	$\varepsilon_t * \varepsilon_{t-1}$	ε_t^2	$\varepsilon_t - \varepsilon_{t-1}$	$(\varepsilon_t - \varepsilon_{t-1})^2$	x_{t-2}
1	2	3	4	5	6	7	8	9	10	11
1	15	-	-	-	-	-	-	-	-	-
2	17	15	16,55	0,55	2,84	-1,5	0,3	-	-	-
3	23	17	18,75	4,25	0,55	2,34	18,06	-3,7	13,68	15
4	27	23	25,37	1,63	4,25	6,93	2,66	2,62	6,86	17
5	32	27	29,78	2,22	1,63	3,62	4,92	-0,59	0,35	23
6	26	32	35,29	-6,29	2,22	-20,6	86,3	11,51	132,48	27
7	21	26	28,68	-7,68	-9,29	71,35	58,98	1,61	2,59	32
8	18	21	23,16	-5,16	-7,68	39,63	26,63	-2,62	6,35	26
9	15	18	19,55	-4,55	-5,16	23,48	20,70	-0,61	0,37	21
10	19	15	16,55	2,45	-4,55	-11,15	6,00	7,0	49	18
11	24	19	20,96	3,04	2,45	7,45	9,24	-0,59	0,35	15
12	33	24	26,47	6,53	3,04	19,85	42,64	-3,49	12,18	19
13	37	33	36,40	0,6	6,53	3,92	0,36	6,47	41,86	24
14	41	37	40,81	0,19	0,6	0,11	0,36	0,41	1,68	33
15	43	41	45,22	-1,78	0,19	-0,34	3,17	1,97	3,88	37
16	45	43	47,43	-2,43	-1,78	4,33	5,9	0,75	0,56	41
17	47	45	49,64	-2,64	-2,43	6,42	6,97	0,21	0,04	43
18	49	47	51,24	-2,84	-2,64	7,5	8,07	0,2	0,04	45
Σ	530					162,28	301,26		272,2	434

Продовження таблиці 4.18

$x_t * x_{t-1}$	$x_t * x_{t-2}$	$x_{t-1} * x_{t-2}$	x_{t-1}^2	x_{t-2}^2	$0,1175 x_{t-1}$	$1,061 x_{t-2}$	\bar{x}_t	$\varepsilon_t = x_t - \bar{x}_t$	ε_t^2	$\varepsilon_t - \varepsilon_{t-1}$	$(\varepsilon_t - \varepsilon_{t-1})^2$
12	13	14	15	16	17	18	19	20	21	22	23
-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-
391	345	255	529	225	1,99	15,91	17,9	5,1	26,01	-	-
621	459	391	729	289	2,70	18,04	20,74	6,26	39,8	-1,13	1,27
864	736	621	1024	529	3,17	24,40	27,57	4,43	19,62	1,83	3,35
832	702	864	676	729	3,76	28,64	32,30	-6,3	39,69	10,73	115,13
546	672	832	441	1024	3,05	33,95	37,00	-16,0	25,6	9,7	94,09
378	468	546	324	676	2,46	27,58	30,04	-12,04	144,96	-3,96	15,68
270	315	378	225	441	2,11	22,28	24,38	-9,38	86,49	-2,66	7,07
285	342	270	361	324	1,76	19,09	20,85	-1,85	3,42	-7,52	56,55
456	360	285	576	225	2,23	15,92	18,15	5,95	35,4	-7,80	60,84
792	627	456	1089	361	2,82	25,46	22,98	10,02	100,4	-4,07	16,54
1221	888	792	1369	576	3,87	26,52	29,33	7,67	58,82	2,35	5,52
1517	1353	1221	1681	1089	4,34	33,20	37,54	3,46	11,97	4,21	17,72
1763	1591	1517	1849	1369	4,82	39,25	44,07	-1,07	1,14	4,53	20,52
1935	1845	1763	2025	1681	5,05	43,50	48,55	-3,55	12,6	-2,48	6,15
2115	1935	1935	2209	1849	5,29	45,62	50,91	-3,91	15,29	0,36	0,12
2303	2205	2115	2405	2025	5,52	47,74	52,26	-4,26	18,14	0,35	0,12
16289	14893	14241	17508	13187				15,47	428,33		

Використовуючи перші 18 членів ряду, складемо одночленну модель:

$$\bar{X}_t = a_1 X_{t-1}.$$

Визначаємо a_1 (табл. 4.18, графи 2, 3):

$$a = \frac{\sum_{l=2}^{18} \bar{X}_t}{\sum_{l=2}^{17} X_{t-1}} = \frac{530}{481} = 1,103. \quad (4.105)$$

Обчислюємо значення $\bar{X}_t = 1,103 X_{t-1}$ і залишків $\varepsilon_t = \bar{X}_t - X_t$ (графи 4, 5). Для використання першого критерію автокорельованості складаємо циклічний ряд ε_{t-1} (графа 6), обчислюємо $\varepsilon_t * \varepsilon_{t-1}$ (графа 7) і ε_t^2 (графа 8).

У результаті одержуємо

$$r_1 = r(\varepsilon_t, \varepsilon_{t-1}) = \frac{\sum_{t=k+1}^n \varepsilon_t * \varepsilon_{t-1}}{\sum_{t=k+1}^n \varepsilon_t^2} = \frac{162,28}{301,36} = 0,528. \quad (4.106)$$

За таблицею 4.18 знаходимо $n_1 = 18 - 1 = 17$ і $r > 0$, маємо $r_{1\%} = 0,475$.

Отже, r_1 потрапляє в критичну область при 1 % рівні значущості, що дає підставу відкинути гіпотезу неавтокорельованості ε_t .

Таким чином, модель не приймається

$$\bar{X}_t = a_1 X_{t-1}.$$

До такого ж висновку приводить і другий критерій Дж. Неймана.

На підставі граф 8, 10 отримаємо:

$$K = \frac{\frac{1}{n-k-1} \sum_{t=k+2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\frac{1}{n-k} \sum_{t=k+1}^n \varepsilon_t^2} = \frac{272,8/16}{301,26/17} = \frac{17,01}{17,72} = 0,959. \quad (4.107)$$

За таблицею 4.18 знаходимо $n_1 = 17$.

$K_{1\%} = 1,035$ – потрапляє в критичну область при 1 % рівні значущості, що дає підставу забракувати гіпотезу відсутності автокорельованості ε_t .

Складаємо двочленну модель

$$\bar{X}_t - \bar{X}_t = a_1 X_{t-1} + a_2 X_{t-2}.$$

Система нормальних рівнянь для визначення параметрів методом найменших квадратів має вигляд:

$$\begin{cases} \sum_{t=3}^{18} X_t * X_{t-1} = a_1 \sum_{t=3}^{18} X_{t-1}^2 + a_2 \sum_{t=3}^{18} X_{t-1} X_{t-2} \\ \sum_{t=3}^{18} X_t * X_{t-2} = a_1 \sum_{t=3}^{18} X_{t-1} X_{t-2} + a_2 \sum_{t=3}^{18} X_{t-2}^2 \end{cases} \quad (4.108)$$

Визначивши суми для вирішення системи (табл. 4.18, графи 12 – 16), отримаємо:

$$16289 = 1728 a_1 + 14241 a_2;$$

$$14853 = 14241 a_1 + 13187 a_2,$$

звідси

$$a_1 = 0,1175;$$

$$a_2 = 1,061.$$

У графах 17 – 19 наведені значення x_t , розраховані по формулі:

$$\bar{X}_t = 0,1175 X_{t-1} + 1,061 X_{t-2}.$$

Відхилення ε_t знаходимо за графою 20 і критерієм Дж. Неймана, перевіряємо неавтокорельованість залишків.

З граф 21 та 23 таблиці 4.18 маємо:

$$K = \frac{\frac{1}{n-k-1} \sum (\varepsilon_t - \varepsilon_{t-1})^2}{\frac{1}{n-k} \sum \varepsilon_t^2} = \frac{528,12/15}{428,33/16} = \frac{35,21}{26,77} = 1,315. \quad (4.109)$$

За таблицею 4.18 маємо:

$$K_{5\%}(16) = 1,309 \text{ при } r > 0;$$

$$K_{5\%}(16) = 2,9577 \text{ при } r < 0.$$

Отже, розрахункове значення K потрапляє в допустиму область при 5 % рівні значущості, що дає підставу для ухвалення гіпотези неавтокорельованості залишків ε_t для затвердження двочленної моделі:

$$\bar{X}_t = 0,1175 X_{t-1} + 1,061 X_{t-2}.$$

При середньоквадратичному відхиленні:

$$\sigma_\varepsilon = \sqrt{\frac{\sum_{t=k+1}^n \varepsilon_t^2}{n-k}} = \sqrt{\frac{428,33}{16}} = 5,17, \quad (4.110)$$

помилка прогнозу

$$B_{cp} = \{ |X_t - \bar{X}_t| \leq t_\alpha * 5,17 \} = P_\alpha;$$

при 90 % гарантійної вірогідності

$$t_\alpha = 1,74,$$

помилка прогнозу не перевищить 8,84.

Прогноз на 19 і 20 періоди

$$X_{19} = 55,61; X_{20} = 58,20$$

з 90 %-ю вірогідністю не переходить за межі:

$$\bar{X} - 8,84 \leq 0,1175 X_{t-1} + 1,061 X_{t-2} \leq \bar{X} + 8,84. \quad (4.111)$$

Таким чином, представлені методики оцінки динаміки економічних процесів і розробки відповідних економетричних моделей динаміки дають змогу приймати управлінські рішення з урахуванням періодів функціонування підприємства та прогнозування економічних показників.

Питання і завдання для самоконтролю до розділу 4

Питання для самоконтролю:

1. Назвіть основні принципи при побудові економетричних моделей.
2. Охарактеризуйте основні критерії оцінки адекватності економетричних моделей.
3. Що таке мультиколінеарність? Назвіть причини її виникнення.
4. У чому полягає парний регресійний аналіз?
5. У чому полягає кількісний регресійний аналіз? Який вигляд має кількісна регресійна модель?
6. Охарактеризуйте етапи побудови багатофакторної економетричної моделі.
7. Охарактеризуйте t-критерій Ст'юдента і F-критерій Фішера для оцінки адекватності багатофакторної економетричної моделі.
8. Охарактеризуйте тест Дарбіна-Уотсона для оцінки адекватності багатофакторної економетричної моделі.
9. Проінтерпретуйте отримані результати на основі розробленої Вами багатофакторної економетричної моделі.
10. Охарактеризуйте узагальнені економетричні моделі.
11. Назвіть види узагальнених економетричних моделей і охарактеризуйте їх.
12. Назвіть основні поняття і визначте сутність динамічних процесів.
13. Що таке часовий ряд і назвіть напрями його оцінки?
14. Що таке авторегресія і як будуються авторегресійні моделі?
15. Назвіть статистичні критерії оцінки автокорельованості залишків і як вони визначаються.

Завдання для самоконтролю:

1. За статистичними даними 10 підприємств (табл. 4.19) розробити рівняння регресії рівня витрат на виробництво продукції (Рввп) від фондоозброєності праці робітників (ФЗп):
 - побудувати поле кореляції і за ним визначити характер та обґрунтувати математичну форму рівняння регресії;
 - визначити коефіцієнти регресії a_0 та a_1 , їх економічний зміст, записати рівняння регресії;
 - визначити коефіцієнти кореляції;
 - визначити з ймовірністю 0,95 довірчі границі помилки апроксимації, записати рівняння регресії в остаточному вигляді;
 - обґрунтувати сутність отриманих результатів.

Таблиця 4.19 – Статистичні дані діяльності підприємств

№ з/п	Рввп, коп./грн.	ФЗп, тис.грн./чол.
1	91,7	1,9
2	91,2	2,1
3	87,7	5,4
4	89,2	2,7
5	90,2	2,2
6	88,8	2,9
7	91,7	2,6
8	92,8	2,7
9	85,7	6,3
10	91,1	4,6

2. За статистичними даними 10 підприємств (табл. 4.20) розробити рівняння регресії рівня витрат на виробництво продукції (Рввп) від фондівіддачі основних засобів (ФФоз):

- побудувати поле кореляції і за ним визначити характер та обґрунтувати математичну форму рівняння регресії;
- визначити коефіцієнти регресії a_0 та a_1 , їх зміст, записати рівняння регресії;
- визначити коефіцієнти кореляції;
- визначити з ймовірністю 0,95 довірчі границі помилки апроксимації, записати рівняння регресії в остаточному вигляді;
- обґрунтування економічної сутності отриманих результатів.

Таблиця 4.20 – Статистичні дані діяльності підприємств

№ з/п	Рввп, коп./грн.	ФФоз, тис. грн./тис. грн.
1	91,7	12,9
2	91,2	12,6
3	87,7	11,9
4	89,2	12,3
5	90,2	12,4
6	88,8	11,6
7	91,7	12,7
8	92,8	12,9
9	85,7	11,2
10	91,1	12,8

3. За статистичними даними 10 підприємств (табл. 4.21) розробити рівняння регресії рентабельності реалізації продукції (R_{rp}) від коефіцієнту вибуття робітників ($K_{вр}$):

- побудувати поле кореляції і за ним визначити характер та обґрунтувати математичну форму рівняння регресії;
- визначити коефіцієнти регресії a_0 та a_1 , їх економічний зміст, записати рівняння регресії;
- визначити коефіцієнти кореляції;
- визначити з ймовірністю 0,95 довірчі границі помилки апроксимації, записати рівняння регресії в остаточному вигляді;
- обґрунтування сутності отриманих результатів.

Таблиця 4.21 – Статистичні дані діяльності підприємств

№ з/п	R_{rp} , коп./грн.	$K_{вр}$, %
1	4,3	4,1
2	4,8	3,3
3	5,3	3,5
4	6,8	2,8
5	5,5	3,9
6	5,2	3,6
7	6,3	2,9
8	7,2	2,7
9	6,4	2,9
10	7,9	2,1

4. Представлені статистичні дані собівартості пасажироперевезень міським електричним транспортом (табл. 4.22), вирівняти за ковзною середньою і побудувати графік і обґрунтувати тенденції зміни собівартості пасажироперевезень.

Таблиця 4.22 – Статистичні дані собівартості пасажироперевезень по депо

t	Собівартість С, коп
1	2
1	89,9
2	90,9
3	87,6
4	87,5
5	88,2
6	89,9

Продовження таблиці 4.22

1	2
7	90,5
8	92,8
9	92,1
10	92,8
11	91,8
12	93,1

5. На основі даних динаміки статистичних показників представлених в таблиці 4.23 виявіть загальну тенденцію їх зміни, побудуйте економетричну модель динаміки і розрахуйте прогноз на 2 наступних роки.

Таблиця 4.23 – Динаміка статистичних показників

Роки	t	Значення показника, С, коп.	Квартал			
			I	II	III	IV
2002	1	68,59	72,4	65,88	69,53	66,54
2003	2	69,95	72,8	67,34	71,21	68,44
2004	3	75,37	74,53	77,49	72,88	76,59
2005	4	75,49	79,22	73,21	73,34	76,19
2006	5	82,34	77,34	82,29	85,45	87,29
2007	6	91,80	88,92	90,72	93,19	94,39

6. Щомісячний пробіг рухомого складу міського електротранспорту на 1000 пасажирів, що перевозяться, представлений в таблиці 4.24. Розрахуйте параметри авторегресійної моделі і складіть прогноз на наступні 2 місяці.

Таблиця 4.24 – Щомісячний пробіг рухомого складу міського електротранспорту на 1000 пасажирів

Період (t)	Щомісячний пробіг рухомого складу міського електротранспорту на 1000 пасажирів (x_t), км.
1	2
1	167,1
2	163,3
3	168,4
4	158,9
5	160,4
6	160,5
7	177,3

1	2
8	171,6
9	173,9
10	174,2
11	174,8
12	171,3
13	169,7
14	170,2
15	173,2
16	185,5

7. Щомісячна реалізація покрівельних матеріалів (в тисячах штук) заводом за 18 місяців представлена в таблиці 4.25. Треба скласти авторегресійну економетричну модель і спрогнозувати місячну потребу в покрівельних матеріалах на 2 наступні місяці.

Таблиця 4.25 – Щомісячна реалізація покрівельних матеріалів заводом

Період (t)	Щомісячна реалізація покрівельних матеріалів (x_t), тис. штук
1	32
2	36
3	33
4	37
5	32
6	29
7	31
8	38
9	35
10	39
11	27
12	33
13	37
14	43
15	46
16	49
17	51
18	54

РОЗДІЛ 5

МЕТОДИ МОДЕЛЮВАННЯ У ЗАДАЧАХ ОПЕРАТИВНОГО ОЦІНЮВАННЯ СТАНІВ СКЛАДНИХ ОБ'ЄКТІВ НЕРУХОМОГО МАЙНА З ВИКОРИСТАННЯМ НЕЧІТКИХ ІНТЕРВАЛЬНИХ УЯВЛЕНЬ

5.1 Розробка методу оперативного оцінювання станів складних об'єктів нерухомого майна з використанням нечітких інтервальних уявлень

Вибір нечітких правил, змінних і їх значень безпосередньо пов'язаний зі знаннями про предметну область, про особливості управління її процесами. Знання можуть бути вилучені з досвіду кваліфікованих спеціалістів, а також із аналізу функціонування аналогів та прототипів.

У зв'язку з цим важливими факторами є побудова та реалізація нечітких алгоритмів, вибір функцій належності і їх дискретизація, визначення методів фаззіфікації та дефаззіфікації [73].

Нехай заданий НДО [80]. Під НДО будемо розуміти взаємодіючі об'єкти і процеси нечітких мережевих моделей, які створюються у процесі моделювання, характеризуються структурою, мають хоча б один нечіткий атрибут, а зміст атрибутів і властивостей змінюється в часі згідно динаміки розвитку змодельованих процесів.

Виділимо етапи, що відображають суть метода.

Етап 1. Нехай НДО описується такими елементами [74 – 76]:

$\tilde{P}_i(in)$	\tilde{t}_i	$\tilde{P}_i(out)$	μ	$\Delta\mu_{i\tilde{A}}$	X	Y	Z	$\tilde{M}_{p_j}(in)$	$\tilde{L}\{x_u\}$	\tilde{M}_{p_j}
-------------------	---------------	--------------------	-------	--------------------------	-----	-----	-----	-----------------------	--------------------	-------------------

Рисунок 5.1 – Елементи, що складають НДО

На рисунку 5.1 $\tilde{P}_i(in)$ – ідентифікація вхідного об'єкта;

\tilde{t}_i – ідентифікація дії (переходу);

$\tilde{P}_i(out)$ – ідентифікація вихідного об'єкта;

μ – значення функції належності вхідного об'єкта;

$\Delta\mu_{i\tilde{A}}$ – значення функції належності, представлене через аналітичну форму деякого інтервального значення ($\Delta\mu_{i\tilde{A}}(x)$), що є принциповим;

X, Y, Z – координати об'єкта в ГІС.

ГІС – системи, призначені для збору, збереження, аналізу і графічної візуалізації просторових даних та пов’язаної з ними інформації про представлені в ГІС об’єкти. Це інструменти, що дозволяють користувачу шукати, аналізувати і редагувати цифрові карти, управляти територіями, а також отримувати додаткову інформацію про об’єкти, наприклад, адреса та телефон людини з врахуванням просторових координат X, Y, Z . Прикладами таких ГІС є програми ArcGIS, ArcPad, NAVSTAR GPS [77];

$\tilde{M}_{p_j}(in)$ – маркування деякої вихідної позиції $\tilde{p}_j \in \{\tilde{P}_i(out)\}$

переходу \tilde{t}_i мережі;

$\tilde{L}\{x_u\}$ – деякий предикат на моделі, який визначається як:

$$\tilde{L}\{x_u\} = \begin{cases} 1, & \text{if } \tilde{L}\{x_u\} = true; \\ 0, & \text{if } \tilde{L}\{x_u\} = false; \end{cases} \quad (5.1)$$

$\tilde{M}_{\tilde{p}_j}$ – вектор поточного маркування.

Еман 2. Нехай є деяка функція $\Delta\mu_{i\tilde{A}}(x)$, яка описується залежністю:

$$y_1(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ \frac{c-x}{c-b}, & b < x \leq c \\ 0, & c < x \end{cases} \quad (5.2)$$

і функція $\Delta\mu_{2\tilde{A}}(x)$, яка описується залежністю:

$$y_2(x) = 1 - e^{-k(x-f)^2}. \quad (5.3)$$

Результуюче значення $\Delta\mu_{i\tilde{B}}(x)$ буде знаходитись в межах «більше, ніж $\Delta\mu_{i\tilde{A}}(x)$ і менше, ніж $\Delta\mu_{2\tilde{A}}(x)$ ». Як показано на прикладах (рис. 5.2 а, рис. 5.2 б), перетин двох функцій належності (5.2) і (5.3), представлених через деякі інтервальні значення, утворює область перетину, площа якої описується деякою фігурою.

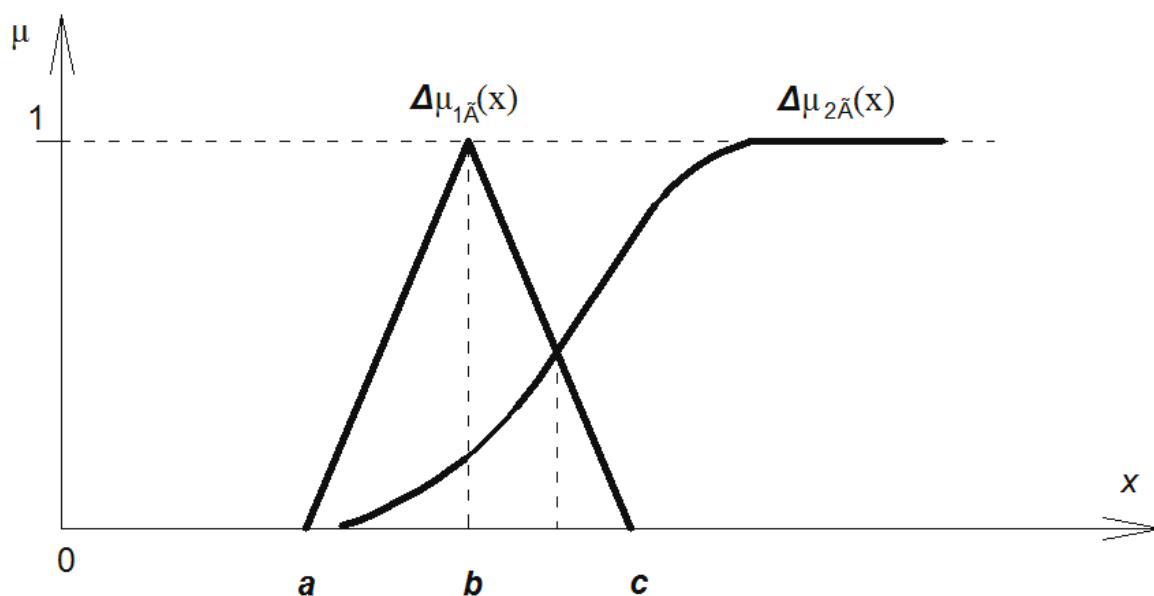


Рисунок 5.2 а – Перетин двох функцій належності, представлених через деякі інтервальні значення

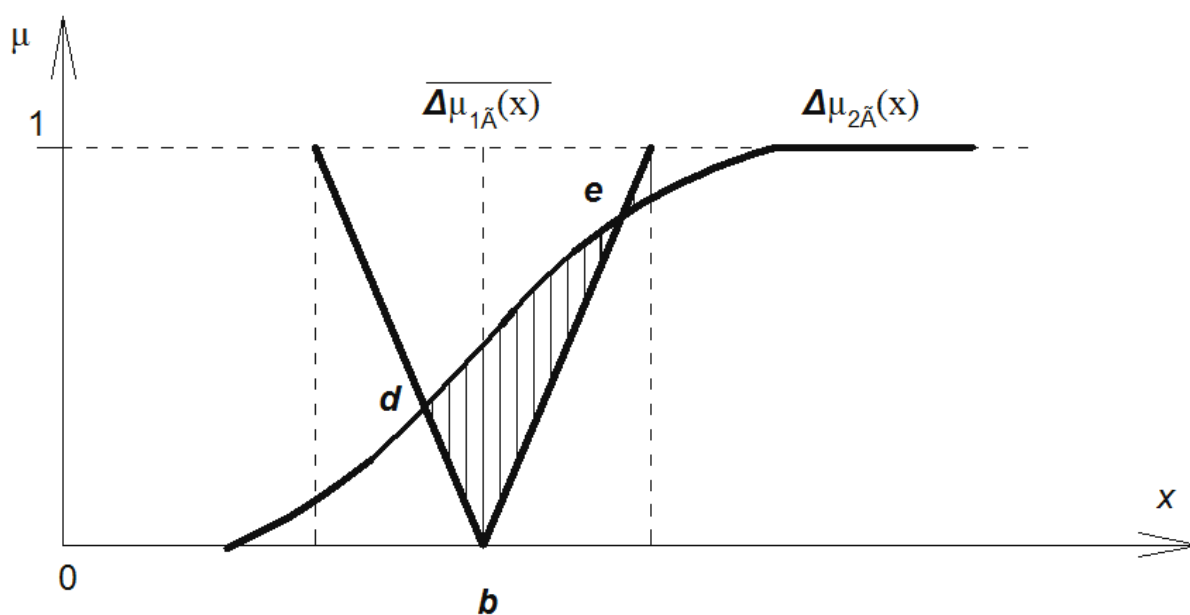


Рисунок 5.2 б – Перетин двох функцій належності (значення $\Delta\mu_{1\tilde{B}}(x)$ визначає штрихова поверхня *deb*)

Етап 3. У даному випадку для визначення значення $\Delta\mu_{1\tilde{B}}(x)$ доцільно та зручно використовувати дефазифікацію на основі принципу «центра ваги (мас)», який дозволяє отримати необхідний результат під час розв’язання широкого класу задач.

$$\Delta\mu_{I\tilde{B}}(x) = \frac{\sum_{i=0}^I \Delta\mu_{i\tilde{A}}(x) S_{i\tilde{A}}}{\sum_{i=0}^I S_{i\tilde{A}}}, \quad (5.4)$$

де $\Delta\mu_{I\tilde{B}}(x)$ – значення результуючої функції належності, представлене через деяке інтервальне значення аналітичної форми;

$\Delta\mu_{i\tilde{A}}(x)$ – значення функції належності, представлене через деяке інтервальне значення аналітичної форми;

$S_{i\tilde{A}}$ – площа області перетину.

Етап 4. Реалізація інтелектуальних обчислювальних механізмів, процедур прямого *modus ponens* (GMP) та зворотного *modus tolens* (GMT) нечіткого логічного виведення [73] (нечітка логіка на інтервалі $0 \leq \Delta\mu_{i\tilde{A}}(x) \leq \Delta\mu_{I\tilde{B}}(x) \leq \Delta\mu_{2\tilde{A}}(x)$) з врахуванням просторових координат об'єкта X, Y, Z .

Етап 4 виконується, якщо X, Y, Z визначено, інакше – останов, виведення повідомлення.

У процедурах *modus ponens* (GMP) *if / then* умови з відомим відношенням $\tilde{R}(x, y)$ та нечітким значенням антецедента \tilde{A}' дозволяють нам визначити наслідок \tilde{B}' :

$$\begin{array}{l} \text{if } x \text{ is } \tilde{A} \text{ then } y \text{ is } \tilde{B} \\ x \text{ is } \tilde{A}' \\ \hline y \text{ is } \tilde{B}' \end{array} \quad (5.5)$$

згідно

$$\tilde{B}' = \tilde{A}' \circ \tilde{R}(x, y) \quad (5.6)$$

з наступною дефазифікацією [73].

Якщо існує деяка множина правил:

$$\begin{array}{l}
\text{if } x \text{ is } \tilde{A}_1 \text{ then } y \text{ is } \tilde{B}_1 \text{ ELSE} \\
\text{if } x \text{ is } \tilde{A}_2 \text{ then } y \text{ is } \tilde{B}_2 \text{ ELSE} \\
\text{if } x \text{ is } \tilde{A}_3 \text{ then } y \text{ is } \tilde{B}_3 \text{ ELSE} \\
\text{.....} \\
\text{if } x \text{ is } \tilde{A}_n \text{ then } y \text{ is } \tilde{B}_n \\
\quad x \text{ is } \tilde{A}' \\
\text{-----} \\
\quad y \text{ is } \tilde{B}'
\end{array} \tag{5.7}$$

то розв’язки для кожного із правил (5.7) знаходяться аналогічно (5.5) з тією відмінню, що тип операції композиції *ELSE* інтерпретується згідно прийнятого оператора побудови нечіткого відношення [73].

Eman 5. У процедурах *modus tolens* (GMT) [73] *if / then* умови з відомим відношенням $\tilde{R}(x,y)$ і нечітким значенням наслідку \tilde{B}' дозволили нам визначити значення антецедента \tilde{A}' :

$$\frac{\begin{array}{l} \text{if } x \text{ is } \tilde{A} \text{ then } y \text{ is } \tilde{B} \\ y \text{ is } \tilde{B}' \end{array}}{\text{-----}} x \text{ is } \tilde{A}' \quad (5.8)$$

ЗГІДНО

$$\tilde{A}' = \tilde{R}(x, y) \circ \tilde{B}' \quad (5.9)$$

з наступною дефаззифікацією [73].

Якщо існує деяка множина правил:

$$\begin{array}{l}
 \text{if } x \text{ is } \tilde{A}_1 \text{ then } y \text{ is } \tilde{B}_1 \text{ ELSE} \\
 \text{if } x \text{ is } \tilde{A}_2 \text{ then } y \text{ is } \tilde{B}_2 \text{ ELSE} \\
 \text{if } x \text{ is } \tilde{A}_3 \text{ then } y \text{ is } \tilde{B}_3 \text{ ELSE} \\
 \dots\dots\dots, \\
 \text{if } x \text{ is } \tilde{A}_n \text{ then } y \text{ is } \tilde{B}_n \\
 \quad y \text{ is } \tilde{B}' \\
 \text{-----} \\
 \quad x \text{ is } \tilde{A}'
 \end{array} \tag{5.10}$$

то розв'язки для кожного із правил (5.10) знаходяться аналогічно (5.8) з тією відміною, що тип операції композиції *ELSE*, по аналогії з GMP, інтерпретується згідно прийнятого оператора побудови нечіткого відношення [73].

Етап 6. Реалізація нечіткого логічного виведення на нечітких алгоритмах та об'єднання кінцевого числа нечітких правил з використанням законів силогізму.

Нехай заданий деякий фрагмент нечіткого алгоритму із таких правил:

$$\begin{array}{l}
 \text{if } x \text{ is } \tilde{A} \text{ then } y \text{ is } \tilde{B} \\
 \text{if } y \text{ is } \tilde{B} \text{ then } z \text{ is } \tilde{C} \\
 \text{if } z \text{ is } \tilde{C} \text{ then } v \text{ is } \tilde{D} \\
 \text{if } v \text{ is } \tilde{D} \text{ then } w \text{ is } \tilde{G} \quad . \\
 \quad x \text{ is } \tilde{A}' \\
 \text{-----} \\
 \quad w \text{ is } \tilde{G}'
 \end{array} \tag{5.11}$$

Із (5.11) ми можемо вивести через правила силогізму таку умову:

$$\text{if } x \text{ is } \tilde{A} \text{ then } w \text{ is } \tilde{G}. \tag{5.12}$$

При відомому відношенні $\tilde{R}_\Sigma(x, w)$ і нечіткому значенні антецедента \tilde{A}' ми можемо визначити наслідок \tilde{G}' у (5.12) і (5.11).

Кожну умову в (5.11) аналітично описано нечіткими відношеннями:

$$\tilde{R}_1(x, y), \tilde{R}_2(y, z), \tilde{R}_3(z, v), \tilde{R}_4(v, w). \quad (5.13)$$

З відношень (5.13) ми можемо вивести нове відношення:

$$\tilde{R}_\Sigma(x, w) = \tilde{R}_1(x, y) \circ \tilde{R}_2(y, z) \circ \tilde{R}_3(z, v) \circ \tilde{R}_4(v, w) \quad (5.14)$$

для умови (5.12). Тоді рішення (5.11) може бути реалізоване як:

$$\tilde{G}' = \tilde{A}' \circ \tilde{R}_\Sigma(x, w). \quad (5.15)$$

Для випадків (5.7) і (5.10) важлива також інтерпретація зв'язки ELSE.

В [73] досліджені деякі важливі підходи до розв'язання цього питання. Нехай заданий деякий фрагмент нечіткого алгоритму із таких правил:

$$\begin{array}{l} \text{if } x \text{ is } \tilde{A} \text{ then } y \text{ is } \tilde{B} \\ \text{if } y \text{ is } \tilde{B} \text{ then } z \text{ is } \tilde{C} \\ \text{if } z \text{ is } \tilde{C} \text{ then } v \text{ is } \tilde{D} \\ \text{if } v \text{ is } \tilde{D} \text{ then } w \text{ is } \tilde{G} \quad , \\ w \text{ is } \tilde{G}' \\ \hline x \text{ is } \tilde{A}' \end{array} \quad (5.16)$$

причому відомий наслідок і необхідно знайти нечітке значення антецедента першого правила із (5.16). З (5.16) ми можемо вивести через силогізм іншу умову (5.12).

При відомому відношенні $\tilde{R}_\Sigma(x, w)$, яке визначене згідно (5.14), і нечіткому значенні наслідку \tilde{G}' ми можемо визначити значення антецедента \tilde{A}' . Тоді отримуємо рішення (5.16) як:

$$\tilde{A}' = \tilde{R}_\Sigma(x, w) \circ \tilde{G}'. \quad (5.17)$$

Етап 7. Процедури нечіткого логічного виведення з наступною дефазифікацією визначимо на основі операторів побудови нечітких відношень [75, 78].

Приймемо за основу формалізацію і інтерпретацію динамічних взаємодіючих нечітких процесів на основі апарата НПКМП. Взаємодіючі процеси можуть бути представлені фрагментами мереж НПКМП (рис. 5.3), для яких характерні різні поєднання потужностей множин вхідних $\{\tilde{p}_i(in)\}$ і вихідних $\{\tilde{p}_i(out)\}$ позицій, потужностей множин вхідних $\{\tilde{t}_j(in)\}$ і вихідних $\{\tilde{t}_j(out)\}$ переходів позиції \tilde{p}_j .

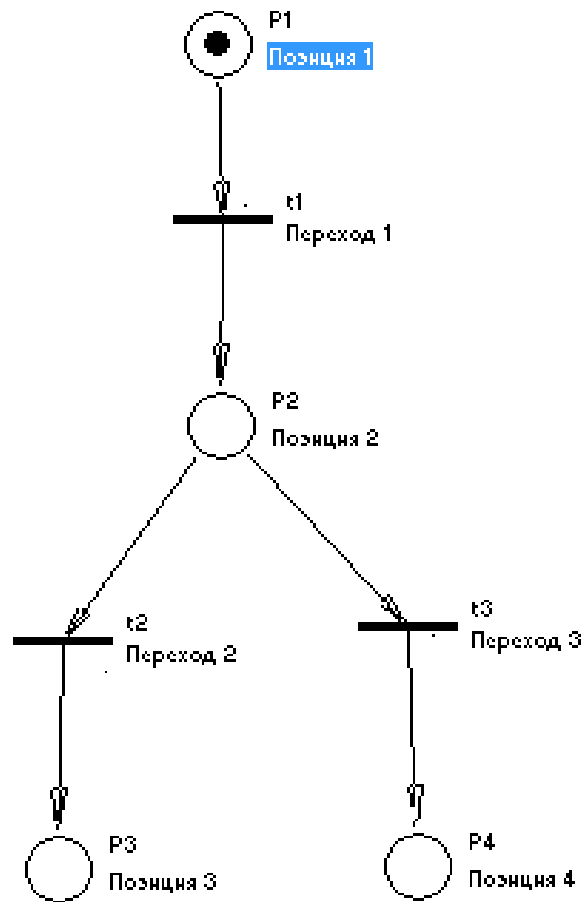


Рисунок 5.3 – Фрагмент моделі

Процедури нечіткого логічного виведення на взаємодіючих процесах можуть бути досліджені на моделях, побудованих композицією фрагментів.

Для фрагмента моделі (рис. 5.3):

$$\exists \tilde{t}_i \in \tilde{T} / (|\{\tilde{p}_i(in)\}| = 1) \text{ and } (|\{\tilde{p}_i(out)\}| = 1) \quad (5.18)$$

запишемо таку конструкцію рішення процедури прямого нечіткого логічного виведення:

$$\begin{array}{l}
 \text{if } \tilde{p}_j \in \{\tilde{p}_i(in)\} \text{ is } \mu_{\tilde{p}_j}(k) \text{ then } \tilde{t}_i \text{ is } \mu_{\tilde{t}_i} \\
 \text{if } \tilde{t}_i \text{ is } \mu_{\tilde{t}_i}(k) \text{ then } \tilde{p}_j \in \{\tilde{p}_i(out)\} \text{ is } \mu_{\tilde{p}_j} \\
 \tilde{p}_j \in \{\tilde{p}_i(in)\} \text{ is } z_{\tilde{p}_j}(k) \\
 \hline
 \tilde{p}_j \in \{\tilde{p}_i(out)\} \text{ is } z_{\tilde{p}_j}(k)
 \end{array} , \quad (5.19)$$

де $z_{\tilde{p}_j}(k)$ у загальному випадку визначене згідно $\mu_{\tilde{p}_j}(k)$

або

$$\begin{array}{l}
 \text{if } \tilde{p}_j \in \{\tilde{p}_i(in)\} \text{ is } \mu_{\tilde{p}_j}(k) \text{ then } \tilde{p}_j \in \{\tilde{p}_i(out)\} \text{ is } \mu_{\tilde{p}_j}(k) \\
 \tilde{p}_j \in \{\tilde{p}_i(in)\} \text{ is } z_{\tilde{p}_j}(k) \\
 \hline
 \tilde{p}_j \in \{\tilde{p}_i(out)\} \text{ is } z_{\tilde{p}_j}(k)
 \end{array} , \quad (5.20)$$

тоді відношення може бути представлено:

$$R(\tilde{p}_j \in \{\tilde{p}_i(in)\}, \tilde{p}_j \in \{\tilde{p}_i(out)\}) = \tilde{R}_1(\tilde{p}_j \in \{\tilde{p}_i(in)\}, \tilde{t}_i) \circ \tilde{R}_2(\tilde{t}_i, \tilde{p}_j \in \{\tilde{p}_i(out)\}). \quad (5.21)$$

З врахуванням (5.20) отримаємо рішення у просторі функцій належності:

$$z_{\tilde{p}_j \in \{\tilde{p}_i(out)\}}(k) = z_{\tilde{p}_j \in \{\tilde{p}_i(in)\}}(k) \wedge \mu_R(\tilde{p}_j \in \{\tilde{p}_i(in)\}, \tilde{p}_j \in \{\tilde{p}_i(out)\}). \quad (5.22)$$

Етап 8. Останов

На основі композиції фрагментів мережі, згідно логіки взаємодії процесів, формуються моделі і обчислюються значення функцій належності очікуваних рішень та їх значень.

Приведемо приклад моделювання процесів з використанням запропонованого метода для фрагмента мережевої моделі (рис. 5.3).

Процес моделювання представимо у вигляді початкового стану фрагмента моделі (рис. 5.4) і подальшого стану фрагмента моделі, після спрацювання дозволеного переходу \tilde{t}_{11} (рис. 5.5).

M_0	11	11	21	$\mu_{\tilde{p}}$	$\Delta\mu_{i\tilde{A}}$	X	Y	Z	0	$\tilde{L}\{x_u\}$	
	11	11	21	$\mu_{\tilde{r}}$	$\Delta\mu_{i\tilde{A}}$	X	Y	Z	0	$\tilde{L}\{x_u\}$	
	21	11	21	$\mu_{\tilde{p}}$	$\Delta\mu_{i\tilde{A}}$	X	Y	Z	0	$\tilde{L}\{x_u\}$	1000

Рисунок 5.4 – Початковий стан фрагмента моделі

M_1	21	21, 22	31, 32	$\mu_{\tilde{p}}$	$\Delta\mu_{i\tilde{A}}$	X	Y	Z	1	$\tilde{L}\{x_u\}$	0100
M_2	21	21	31	$\mu_{\tilde{p}}$	$\Delta\mu_{i\tilde{A}}$	X	Y	Z	1	$\tilde{L}\{x_u\}$	0010
M_3	21	22	32	$\mu_{\tilde{p}}$	$\Delta\mu_{i\tilde{A}}$	X	Y	Z	1	$\tilde{L}\{x_u\}$	0001

Рисунок 5.5 – Наступні стани фрагмента моделі

У результаті виконання процесів формуються множини маркованих позицій ($\{\tilde{p}_{11}\}$, $\{\tilde{p}_{21}\}$, $\{\tilde{p}_{31}\}$, $\{\tilde{p}_{32}\}$) і множина виконаних переходів ($\{\tilde{t}_{11}\}$, $\{\tilde{t}_{21}\}$, $\{\tilde{t}_{22}\}$). Значення функцій належності $\mu_{\tilde{p}}$, $\mu_{\tilde{r}}$, $\Delta\mu_{i\tilde{A}}$, координатні значення X, Y, Z , умова виконання предиката $\tilde{L}\{x_u\}$ визначаються для позицій і переходів згідно приведеної вище стратегії (аналогічно 5.19 – 5.22).

Нечітке логічне виведення на основі нечіткої логіки. Застосування моделі дає можливість дослідити динаміку розвитку процесів і їх взаємодію у нечіткому просторі станів на основі положень нечіткої логіки, що підвищує достовірність рішень в умовах апріорної невизначеності.

У зв'язку з цим, для розв'язання практичних задач згідно (5.11) – (5.15) доцільно реалізувати таку стратегію [76]:

1. Обчислюємо функцію належності:

$$\mu^*(f_n), \quad (5.23)$$

задану на деякому інтервалі нечіткої логіки.

2. Для обчисленої функції (5.23) методами дефазифікації [74] знаходимо значення аргументу k_0 цієї функції.

3. Для обчисленого значення аргументу k_0 знаходимо значення вектора:

$$\mu^*(f_n) /_{k=k_0}. \quad (5.24)$$

4. Функція (5.23) визначає функцію антецедента першого правила в (5.11) для реалізації відношення (5.14).

5. Значення вектора (5.24) визначає вхідне значення \tilde{A}' в правилах (5.15).

6. Згідно (5.15) обчислюємо вектор консеквента.

7. Результат дефаззифікації вектора консеквента визначає потрібний аргумент і значення функції належності консеквента $\mu_w(k)'$.

З врахуванням запропонованої стратегії, замінюючи значення лінгвістичних змінних у (5.11) – (5.15) функціями належності $\mu(k)$, отримаємо:

$$\mu_w(k)' = \mu^*(f_n) /_{k=k_0} \circ \mu(x, w). \quad (5.25)$$

З метою дослідження динаміки моделюючих процесів на практиці важливо реалізувати процедури моделювання на нечітких мережах Петрі або їх фрагментах під час подання значення вхідного вектора $\mu_{\tilde{p}_j}(k)$ і значення $z_{\tilde{p}_j}(k)$ для всіх:

$$\tilde{p}_j \in \{\tilde{p}_i(in)\} \quad (5.26)$$

на основі нечіткої інтервальної логіки. Згідно викладеної вище стратегії, при нечітких вхідних векторах, представлених через деякі інтервальні значення, обчислимо для всіх (5.26):

$$\mu_{\tilde{p}_j}(k) = \mu^*(f_n) \quad (5.27)$$

і

$$z_{\tilde{p}_j}(k) = \mu^*(f_n) /_{k=k_0}. \quad (5.28)$$

Тоді рішення виду (5.22) може бути подане як:

$$(z_{\tilde{p}_j \in \{\tilde{p}_i(out)\}}(k)) = ((z_{\tilde{p}_j(k)} = \mu^*(f_n) /_{k=k_0}) / z_{\tilde{p}_j \in \{\tilde{p}_i(in)\}}(k)) \wedge \quad (5.29)$$

$$\wedge (\mu_R(\tilde{p}_j \in \{\tilde{p}_i(in)\}, \tilde{p}_j \in \{\tilde{p}_i(out)\})).$$

Аналогічно (5.29) можемо отримати рішення для фрагментів з іншими поєднаннями вхідних і вихідних вершин. На основі (5.29), шляхом композиції деяких фрагментів мережі, здійснюємо моделювання процесів предметної області, що і є потрібним рішенням.

Визначимо шляхи побудови нечітких відношень типу $\tilde{R}(x, y)$, $\tilde{R}_\Sigma(x, w)$ і інтерпретації операції композиції *ELSE* [73] у сформульованих нами рішеннях (5.7) і (5.10). У [73] розглянуті і узагальнені деякі із існуючих рішень. У [73, 75] приведені оператори знаходження функцій належності відношень $\mu(x, y)$ і інтерпретація операції композиції *ELSE* [74] у (5.7) і (5.10).

Нечіткі оператори значення [74, 75] знайшли застосування у технічних додатках систем управління і контролю на основі нечіткого логічного виведення. Додатково до розглянутих у [74, 75] операторів визначимо і уточнимо можливі рішення і знаходження відношень у задачах композиції фрагментів алгоритмів [74].

Рішення на основі сукупності нечітких *if / then* умов у нечітких алгоритмах математично фактично визначаються знаходженням відношень $\tilde{R}(x, y)$, $\tilde{R}_\Sigma(x, w)$ і, власне, їм еквівалентні в сукупності з вектором входу.

Перейшовши до функцій належності у поданні правил продукції типу (5.5), нечітке відношення $\tilde{R}(x, y)$ у рішенні (5.5) ми можемо представити для випадку неперервних функцій належності таким чином:

$$\tilde{R}(x, y) = \int_{(x, y)} \mu(x, y) / (x, y), \quad (5.30)$$

а також для випадку дискретних функцій належності:

$$\tilde{R}(x, y) = \sum_{(x, y)} \mu(x, y) / (x, y), \quad (5.31)$$

де $\mu(x, y)$ – функція належності відношення.

Слід відмітити, що дискретний ряд алгоритмічно більш кращий і знаходження (5.31) є більш прийнятним у практичних реалізаціях у порівнянні з (5.30). Тому, не йдучи від відомих узагальнень, у подальшому ми будемо орієнтуватися на рішення (5.31).

Процедури обчислення відношень шляхом знаходження максимальних компонент на множині мінімальних значень функцій належності [75].

Склад з максимальним мінімумом двох нечітких відношень, який був використаний для композиції правил у приведених вище процедурах, використовує оператори нечітких множин:

- оператор знаходження максимуму \vee ;
- оператор знаходження мінімуму \wedge .

Нехай є два нечітких відношення:

$$\tilde{R}(x, y), \tilde{R}(y, z), \quad (5.32)$$

відношення визначені на декартових добутках:

$$X \times Y, Y \times Z. \quad (5.33)$$

У випадку дискретної декартової площини:

$$X \times Z \quad (5.34)$$

знаходження оператора композиції фрагментів алгоритмів і моделей представимо як:

$$\tilde{R}_\Sigma = \sum_{X \times Z} \vee_y (\mu_{\tilde{R}_1(x, y)} \wedge \mu_{\tilde{R}_2(y, z)}) / (x, z). \quad (5.35)$$

З рівняння (5.35) знаходимо, що ступінь належності кожної (x, z) у результуючому відношенні \tilde{R}_Σ виглядає таким чином:

$$\mu_{\tilde{R}_1 \circ \tilde{R}_2}(x, z) = \vee_y (\mu_{\tilde{R}_1(x, y)} \wedge \mu_{\tilde{R}_2(y, z)}), \quad (5.36)$$

де зовнішній максимум прийнятий відносно елементів у загальній границі.

Очевидно, що (5.35) і (5.36) природним чином можуть бути розповсюджені на будь-яке кінцеве число правил продукції.

Склад з максимальним мінімумом використовується зазвичай у задачах оперативного оцінювання станів об'єктів дослідження і прикладних програмах управління на основі нечіткої логіки.

Процедури обчислення відношень шляхом знаходження максимального — складу на множині функцій належності [75].* Особливість даного підходу і відповідних рішень заключається у тому, що ми можемо використовувати множення, додавання або деяку іншу двійкову операцію (*) на місці мінімуму (\wedge) у рівняннях (5.35) і (5.36) при умові виконання операції максимуму відносно змінної y .

Цей тип складу двох нечітких відношень відомий як «максимально –*». З врахуванням викладеного вище матеріалу визначимо відповідні залежності.

Максимально –* склад для дискретної декартової площини (5.34) може бути представлений:

$$\tilde{R}_{\Sigma} = \sum_{X \times Z} \bigvee_y (\mu_{\tilde{R}_1(x, y)} * \mu_{\tilde{R}_2(y, z)}) / (x, z). \quad (5.37)$$

З рівняння (5.37) бачимо, що ступінь належності кожної (x, z) у результуючому відношенні \tilde{R}_{Σ} виглядає таким чином:

$$\mu_{\tilde{R}_1 \circ \tilde{R}_2}(x, z) = \bigvee_y (\mu_{\tilde{R}_1(x, y)} * \mu_{\tilde{R}_2(y, z)}), \quad (5.38)$$

де зовнішній максимум прийнятий відносно елементів y загальної границі.

Очевидно, що (5.37) і (5.38) природним чином можуть бути розповсюджені на будь-яке кінцеве число правил продукції. Розглянемо деякі приватні випадки застосування максимально –* складу композиції.

Процедури обчислення відношень шляхом знаходження максимально-середнього складу на множині функцій належності [75]. У процедурах з максимально-середнім складом використовуємо такі найбільш очевидні підходи:

1. Застосовуємо операцію знаходження арифметичного середнього замість операції (*) у рівнянні (5.37) і (5.38).

У випадку дискретної декартової площини (5.34) відношення може бути представлено:

$$\tilde{R}_{\Sigma} = \sum_{X \times Z} \bigvee_y \left(\frac{1}{2} (\mu_{\tilde{R}_1(x, y)} + \mu_{\tilde{R}_2(y, z)}) \right) / (x, z). \quad (5.39)$$

З рівняння (5.39) бачимо, що ступінь належності кожної (x, z) у результуючому відношенні \tilde{R}_{Σ} виглядає таким чином:

$$\mu_{\tilde{R}_1 \circ \tilde{R}_2}(x, z) = \bigvee_y \left(\frac{1}{2} (\mu_{\tilde{R}_1(x, y)} + \mu_{\tilde{R}_2(y, z)}) \right), \quad (5.40)$$

де зовнішній максимум прийнятий відносно елементів y загальної границі.

Очевидно, що (5.39) і (5.40) природним чином можуть бути розповсюджені на будь-яке кінцеве число правил продукції.

2. Застосовуємо операцію знаходження квадратичного середнього замість операції (*) у рівняннях (5.37) і (5.38). З врахуванням викладеного вище можемо визначити відповідні залежності.

У випадку дискретної декартової площини (5.34) вираз (5.37) може бути представлений:

$$\tilde{R}_{\Sigma} = \sum_{X \times Z} \vee_y \left(\left(\frac{1}{2} (\mu_{\tilde{R}_1(x, y)}^2 + \mu_{\tilde{R}_2(y, z)}^2) \right)^{\frac{1}{2}} \right) / (x, z). \quad (5.41)$$

З рівняння (5.41) знаходимо, що ступінь належності кожної (x, z) у результуючому відношенні \tilde{R}_{Σ} виглядає таким чином:

$$\mu_{\tilde{R}_1 \circ \tilde{R}_2}(x, z) = \vee_y \left(\left(\frac{1}{2} (\mu_{\tilde{R}_1(x, y)}^2 + \mu_{\tilde{R}_2(y, z)}^2) \right)^{\frac{1}{2}} \right), \quad (5.42)$$

де зовнішній максимум прийнятий відносно елементів у загальній границі.

Очевидно, що (5.41) і (5.42) природним чином можуть бути розповсюджені на будь-яке кінцеве число правил продукції.

3. Застосовуємо операцію знаходження геометричного середнього замість операції (*) у рівняннях (5.37) і (5.38).

У випадку дискретної декартової площини (5.34) вираз (5.37) може бути представлений:

$$\tilde{R}_{\Sigma} = \sum_{X \times Z} \vee_y \left(\left(\mu_{\tilde{R}_1(x, y)} + \mu_{\tilde{R}_2(y, z)} \right)^{\frac{1}{2}} \right) / (x, z). \quad (5.43)$$

З рівнянь (5.43) знаходимо, що ступінь належності кожної (x, z) у результуючому відношенні \tilde{R}_{Σ} виглядає таким чином:

$$\mu_{\tilde{R}_1 \circ \tilde{R}_2}(x, z) = \vee_y \left(\left(\mu_{\tilde{R}_1(x, y)} + \mu_{\tilde{R}_2(y, z)} \right)^{\frac{1}{2}} \right), \quad (5.44)$$

де зовнішній максимум прийнятий відносно елементів у загальній границі.

Очевидно, що (5.43) і (5.44) природним чином можуть бути розповсюджені на будь-яке кінцеве число правил продукції.

Запропоновані підходи знаходження результируючих відношень \tilde{R}_Z можуть бути застосовані в інтелектуальних обчислювальних нечітких алгоритмах практичних додатків до моделювання процедур нечіткого логічного виведення на їх основі. Вибір конкретного підходу із (5.35) – (5.44) визначається особливостями предметної області.

5.2 Формалізація і розробка розширеного методу дихотомії для задачі настроювання параметрів функцій належності

У нечіткій логіці значення будь-якої величини представляються не числами, а словами природної мови, які називаються термами [79]. Для лінгвістичної змінної «дистанція» термами є «далеко» і «близько». Під час реалізації лінгвістичної змінної необхідно дати точний фізичний опис її термів.

Слід відмітити, що визначення ступені належності можливе тільки під час спільної роботи з експертами. Виникає питання про необхідну кількість термів у змінній для достатньо точного представлення фізичної величини. Існує думка, що для більшості додатків достатньо 3 – 7 термів на кожну змінну.

Належність кожного точного значення до одного із термів лінгвістичної змінної визначається за допомогою функції належності [79].

На рисунку 5.6 [80] запропоновані стандартні функції належності, які застосовуються для вирішення більшості задач. Під час розв’язання специфічних задач можна вибрати і інші, більш прийнятні види функцій належності і, як показує практика, досягти кращих результатів роботи системи, ніж під час використання функцій стандартного виду.



Рисунок 5.6 – Основні типи функцій належності

Функції належності на рисунку 5.6 представлені у вигляді трикутника.

У загальному випадку всі функції належності можна розділити на три значення: функції «великої», «малої» і «середньої» величини.

Припустимо, що дана деяка функція належності. Експерт задає її вигляд $\mu_{i\tilde{A}}(x)$ з деякими параметрами k . У процесі нечіткого логічного виведення Заде-Мамдані [73] встановлено, що початкова функція (функції) належності має похибку Δ , яка перевищує деяку допустиму похибку ε . Враховуючи дану ситуацію, що виникає під час розв'язання широкого класу задач:

- на множині функцій $\mu_{i\tilde{A}}(x)$ необхідно запропонувати метод, який мінімізує Δ з врахуванням допустимої похибки:

$$\varepsilon \leq \varepsilon^*, \quad (5.45)$$

де ε^* – гранично допустиме значення помилки;

- визначити і обґрунтувати складність розв'язання поставленої задачі за часовим критерієм τ .

У процесі побудови функції належності виникає необхідність визначення її параметрів і приведення у відповідні з існуючими експертними оцінками функції.

Були розглянуті переваги і недоліки градієнтного методу [81, 82], генетичних алгоритмів [82 – 84], метода бінарного пошуку [82, 85] стосовно даної проблеми, що показали їх деяку функціональну обмеженість.

Одна із основних проблем [82] застосування градієнтного метода пошуку заключається у виборі величини кожного дискретного кроку. Кроки можуть бути постійними або змінними. Другий варіант у реалізації алгоритму більш складний, але зазвичай потребує меншої кількості ітерацій.

Розглянутий алгоритм застосовують тільки для нелінійних функцій. Якщо функція лінійна, з наявністю локальних і глобальних екстремумів, то вибір оптимального значення параметрів ускладнений через можливість пропуску глобальних екстремумів.

Таким чином, недоліком цих алгоритмів [82] є їх обчислювальна ресурсоемність і ускладненість пошуку глобального екстремуму при наявності функції з множиною локальних екстремумів.

Генетичний алгоритм [82 – 84] – це евристичний алгоритм пошуку, що використовується для рішення задач оптимізації і моделювання. Він здійснюється шляхом послідовного підбору, комбінування і варіації потрібних параметрів з використанням механізмів, що нагадують біологічну еволюцію, і є

різновидом еволюційних обчислень. Відмінною особливістю генетичних алгоритмів є акцент на використання оператора «схрещування», який проводить операцію рекомбінації рішень-кандидатів, роль якої аналогічна ролі схрещування у живій природі.

Слід відмітити, що генетичні алгоритми дозволяють з достатньо високою точністю апроксимувати функцію, але вони володіють і рядом суттєвих недоліків, до яких відносять [82]:

- високу обчислювальну складність (що нераціонально при машинній обробці);
- складність у реалізації самого алгоритму;
- нестійкість рішень.

Розробка розширеного удосконаленого методу дихотомії для задачі настроювання параметрів функцій належності [82]. Для розв'язання вище поставленої задачі було запропоновано використання бінарного підходу [86], зокрема – методу дихотомії [85], для підбору параметрів нечітких функцій належності.

Дихотомія (метод половинного ділення) – послідовний поділ на дві частини, що не пов'язані між собою. Дихотомічний поділ у математиці, філософії, логіці і лінгвістиці є способом створення взаємовиключних підрозділів одного поняття або терміна, а також необхідним для створення класифікації елементів [85].

Однією із переваг дихотомічного поділу є зручність його реалізації. Достатньо двох класів, які вичерпують обсяг діленого поняття. Таким чином, дихотомічний поділ завжди пропорційний; члени ділення вичерпують один одного, так як кожний об'єкт діленої множини попадає тільки в один із класів (a або не a); ділення проводиться за однією основою (наявність або відсутність деякої ознаки). Позначивши ділене поняття буквою a і виділивши в його об'ємі деякий вид b , можна розділити об'єм a на дві частини b і не b .

Метод дихотомії у деякій ступені схожий з методом двійкового пошуку, однак відрізняється від нього критерієм «відкидання кінців» [82, 85].

Покажемо поетапну роботу розширеного методу дихотомії стосовно до задачі настроювання параметрів функцій належності.

Етап 1. Нехай задане початкове значення параметрів функцій належності:

$$f(x): [a,b] \rightarrow R, f(x) \in C([a,b]). \quad (5.46)$$

Етап 2. Розіб'ємо заданий інтервал пополам і візьмемо дві симетричні відносно центру точки x_1 і x_2 так, що:

$$x_1 = \frac{a+b}{2} - \delta; \quad x_2 = \frac{a+b}{2} + \delta, \quad (5.47)$$

де δ – деяке число в інтервалі $\left(0, \frac{b-a}{2}\right)$.

Етап 3. Відкинемо той із кінців початкового інтервалу, до якого ближче виявилась одна із двох знову поставлених точок з максимальним значенням (у випадку пошуку мінімуму), якщо:

$$f(x_1) > f(x_2), \quad (5.48)$$

то береться відрізок $[x_1, b]$, а відрізок $[a, x_1]$ відкидається.

Інакше береться дзеркальний відносно середини відрізок $[a, x_2]$, а відкидається $[x_2, b]$.

Етап 4. Процедури етапу 3, обчислювальні процедури дихотомії і процедури нечіткого логічного виведення на деяких із функцій (5.46) виконуються, доки не буде досягнута задана точність ε [82] на множині значень $\{k_i\}$, $i \in I$:

$$|y_{\text{факт}} - y_{\text{очік}}| \leq \varepsilon, \quad (5.49)$$

де ε – числове значення точності апроксимації, тобто гранично допустиме відхилення фактичного значення від очікуваного.

Етап 5. Визначення значень $\{k_i\}$, $i \in I$.

Етап 6. Повторне виконання етапів 2 – 5.

Етап 7. Останов.

Так як на кожній ітерації приходить ся обчислювати нові точки, необхідно запропонувати метод, який дозволить вираховувати лише одну нову точку на черговій ітерації, що суттєво оптимізує процедуру. Такий підхід досягається шляхом дзеркального поділу відрізка в перерізі, тоді в даному випадку запропонований метод можна розглядати як поліпшення існуючих підходів [82].

Слід відмітити також, що дихотомія має властивість збіжності для будь-яких неперервних функцій, у тому числі і недиференційованих [85].

До підбору параметрів нечіткої функції належності метод дихотомії застосовують таким чином [82]:

- під час дефазифікації виділяються деякі значення нечітких функцій належності, що входять у продукцію «Якщо X , то Y ». Аргументом вхідної функції належності подається деяка величина x_{ex} , яка, у відповідності з одним із методів дефазифікації (наприклад, центроїдним), проектується на вихідну функцію Y , в якості виходу отримуємо деяке значення $y_{факт}$;

- суть підбору параметрів полягає у тому, щоб змінити вид вихідної функції належності у відповідність з отриманим від експерта очікуваним значенням $y_{очік}$. У даному випадку має місце процес апроксимації вихідної функції належності з виконанням умови (5.49) [82];

- так як функція належності будь-якого типу має деяку параметричну величину k , то саме вона підлягає підбору значень.

Наприклад, для функції «великої» величини має місце вираз:

$$y = 1 - e^{k(x-a)^2}. \quad (5.50)$$

Параметр k підлягає дихотомії з використанням кроку апроксимації рівного $\frac{\varepsilon}{2}$. Дана величина кроку обумовлена необхідністю виконання виразу (5.49). На виході метод надає знайдені значення параметрів функцій належності.

Дослідження складності методу [82] показало, що кількість ітерацій метода, а відповідно і час його роботи τ мають залежність близьку до комбінаторного пошуку, тобто до експоненціальної складності з врахуванням величини значення точності ε .

Дійсно, пошук на одній ітерації має затрати $\approx n$ операцій. Стратегія дихотомії припускає процедури поділу на два під час реалізації деякої ітерації. Ділення цілих ненульових n -розрядних (не рахуючи знакових розрядів) чисел A і B , представлених в прямому (для простоти) коді, призводить до отримання цілої частки C і цілого залишку 0 , якому присвоюється знак діленого, а знак частки обчислюється як сума по модулю 2 операндів A і B .

Ділення виконується у такій послідовності [86]:

- а) дільник B зсувається вліво (нормалізується) так, щоб у старшому інформаційному розряді опинилася 1, далі підраховується кількість зсувів S . Частка від ділення може бути не більше $S + 1$ розрядів, не рівних нулю;

б) виконується $S + 1$ цикл ділення модулів $|A|$ на $|B'|$, де $|B'|$ – нормалізоване B . У результаті даної операції знаходиться $(S + 1)$ розряд частки, починаючи зі старшого із $(S + 1)$ молодших;

в) отриманий у останньому циклі ділення залишок R_{S+1} зсувається вправо на S розрядів, якщо він позитивний. У випадку отримання від'ємного залишку $R_{S+1} < 0$ до нього для відновлення додається $|B'|$, тобто:

$$[R_{S+1}]_{\text{vost}} = R_{S+1} + |B'|. \quad (5.51)$$

Після цього виконується зсув вправо на S розрядів. У результаті виходить цілий залишок від ділення.

Для визначення верхньої оцінки складності розширеного метода дихотомії при n операціях на кожній ітерації представимо стратегію пошуку як:

$$n + n(n-1) + n(n-1)(n-2) + n(n-1)(n-2)(n-3) + \dots \quad (5.52)$$

Приймаємо до уваги розкладання у ряд показників функції [85] як

$$e^x \approx 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad (5.53)$$

Тоді в області точного настроювання параметрів функцій належності у задачах нечіткого логічного виведення [73] ми можемо визначити верхню оцінку складності, яка не перевищує такої залежності:

$$O(n) \approx ke^n, \quad (5.54)$$

де $O(n)$ – складність метода;

n – кількість обчислювальних ітерацій;

k – параметр, що визначається обчислювальними ресурсами комп'ютера, У даному випадку k деяка нелінійна зростаюча функція.

Очевидно, що залежність (5.54) багато в чому корелюється параметром k , що потребує, можливо, вирішення оптимізаційних задач щодо вибору конфігурації апаратних засобів комп'ютера. Тоді розв'язання задачі може бути представлено як:

$$\begin{aligned} \varepsilon &\rightarrow \min; \\ V &\geq V^*; \\ \varepsilon &\leq \varepsilon^*, \end{aligned} \quad (5.55)$$

де V^* – деяке наперед задане граничне значення швидкості.

Розв'язання задачі (5.55) є нетривіальним, так як залежності ε і V – нелінійні. Враховуючи, що у (5.54)

$$n \approx \frac{I}{\varepsilon},$$

то розв'язання (5.55) одночасно знижує складність (5.54), що важливо у практичних додатках.

Для зменшення складності може бути запропоновано використання метода гілок та границь [85], який застосовують для знаходження оптимальних рішень різних задач оптимізації, особливо дискретної і комбінаторної оптимізації. По суті, метод є варіацією повного перебору з відсівом підмножин допустимих рішень, що завідомо не містять оптимальних рішень. Це призводить до значного зниження складності (5.54).

Робота метода закінчується при досягненні $\varepsilon \leq \varepsilon^*$.

Загальні положення метода можуть бути описані на прикладі пошуку мінімуму або максимуму функції $f(x)$ на множині допустимих значень x . Функція f і x можуть бути довільної природи. Для метода гілок та границь необхідні дві процедури: розгалуження і знаходження оцінок (границь).

Процедура розгалуження складається у розбитті області допустимих рішень на підобласті менших розмірів. Процедuru можна рекурсивно застосовувати до підобластей. Отримані підобласті утворюють дерево, яке називається деревом пошуку або деревом гілок та границь. Вузлами цього дерева є побудовані підобласті [85].

Процедура знаходження оцінок полягає у пошуку верхніх та нижніх границь для оптимального значення на підобласті допустимих рішень.

В основі метода гілок та границь лежить така концепція (наприклад, для задачі мінімізації): якщо нижня границя для підобласті A дерева пошуку більша, ніж верхня границя якої-небудь раніше розглянутої підобласті B , то A може бути виключена із подальшого розгляду (правило відсіву). Мінімальну із отриманих верхніх оцінок запишемо у глобальну змінну m , а будь-який вузол дерева пошуку, нижня границя якого більша значення m , може бути виключеним із подальшого розгляду.

Якщо нижня границя вузла дерева співпадає з верхньою границею, то це значення є мінімумом функції і досягається на відповідній підобласті.

Для оцінки складності у даній роботі було проведене нормування значень точності та кількості операцій до одиничної норми [82], тобто, якщо позначити точність ε' , то: $\varepsilon' = 1 - \varepsilon$.

Кількість операцій обчислюється множенням реального числа виконуваних операторів на коефіцієнт пропорційності a :

$$a = \frac{1}{N}, \quad (5.56)$$

де N – кількість операцій при:

$$\varepsilon' \geq 1. \quad (5.57)$$

Був виконаний вимір часу роботи програми на апаратній платформі комп'ютера з процесором Intel Core2Duo, тактовою частотою 2 ГГц, об'ємом ОЗП 2 Гб і операційною системою Windows XP Professional.

При значенні точності $\varepsilon = 0.01$, час обчислювальної роботи додатка склав: $\tau = 12$ мс; при значенні точності $\varepsilon = 0.001$, час обчислювальної роботи додатка склав: $\tau = 786$ мс.

5.3 Програмна реалізація розробки

Розроблений додаток [82] призначений для реалізації автоматичного настроювання параметрів функцій належності k за деякими вхідними умовами. Користувач має можливість визначати наступні умови: типи вхідної та вихідної фаззи-величин («мала», «середня», «велика»); значення аргументу для вхідної величини $x_{\text{вх}}$ (точка відрахунку дефаззифікації); очікуване значення аргументу вихідної величини $y_{\text{очік}}$ на основі деяких отриманих раніше експертних оцінок; необхідну точність обчислень ε .

Суть проведених обчислень заключається у підборі значень коефіцієнта k , що визначає вид функції вихідної нечіткої величини заданого типу з метою знаходження наближеного значення $y_{\text{факт}}$ до значення $y_{\text{очік}}$ з дотриманням вимог точності обчислень.

Додаток реалізує метод настроювання параметрів під час дефаззифікації нечітких величин [82]:

- визначення типів вхідної та вихідної нечітких величин;
- зазначення точки відрахунку у вигляді аргументу функції вхідної нечіткої величини $x_{\text{вх}}$;

- знаходження значення функції вхідної нечіткої величини, що відповідає заданій точці відрахунку і співставлення значенню вихідної нечіткої величини;
- формування складової підінтегральної фігури, обмеженої функцією вихідної нечіткої величини і лінією проектування значення функції вхідної нечіткої величини на графік функції вихідної нечіткої величини;
- знаходження абсциси центра ваги отриманої фігури, яка визначає $y_{факт}$.

Алгоритм підбору значення коефіцієнта k для вихідної нечіткої величини зводиться до мінімізації (з раніше визначеною точністю ε) абсолютної різниці очікуваного і фактичного значення аргументу [82]:

$$\Delta x = |y_{очік} - y_{факт}|. \quad (5.58)$$

Програмна реалізація даного метода виконана в об'єктно-орієнтованому програмному середовищі C++ Builder [87] і передбачає наявність деяких апроксимуючих припущень, що дозволяють реалізувати обчислення числовими методами [88] за допомогою розбиття аналізованої підінтегральної фігури на ряд простих складових (прямокутні трапеції) з метою знаходження центра ваги. Кількість фігур, що розбиваються, визначається заданою точністю обчислень. До апроксимації можна віднести і механізм підбору коефіцієнта вихідної функції k , який здійснюється методом дихотомії (половинного ділення).

Класова ієрархія програмної реалізації складається із трьох базових класів (рис. 5.7) [82, 87]:

- клас головної інтерфейсної форми TMainForm;
- клас нечіткої величини (опис функції належності) FuzzyValue;
- клас довільного багатокутника MyPoly, що використовується під час знаходження центра ваги у процесі дефаззифікації.

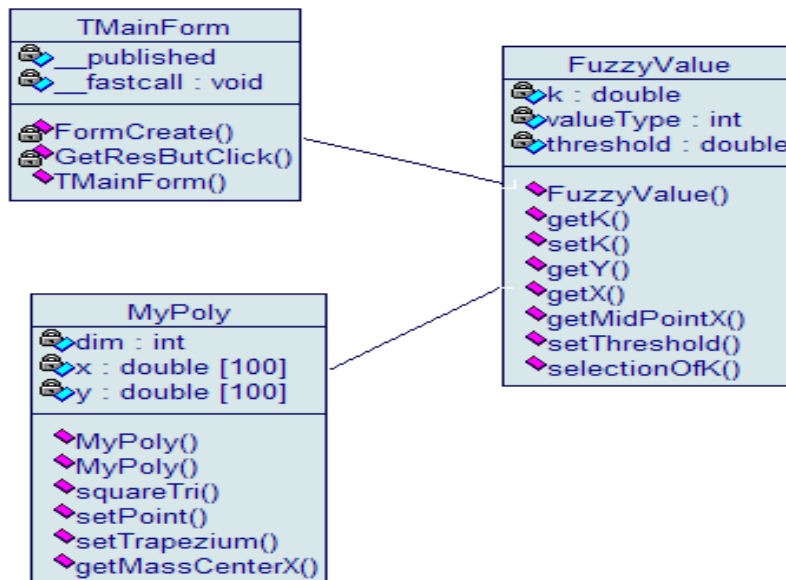


Рисунок 5.7 – Діаграма класів програми

Між всіма класами програми встановлене відношення агрегації. Основна обчислювальна складова визначається у функціях-членах класів.

Основними функціями для реалізації запропонованого метода є функція знаходження центра ваги `getMidPointX()` і функція дихотомічного підбору параметра k `selectionOfK()` (рис. 5.7).

Інтерфейсна частина програми гранично проста і орієнтована на користувача. Вона надає можливість задання для вхідних параметрів: тип нечіткої функції, її параметр k і вхідне значення аргументу; для вихідних параметрів: тип нечіткої функції, очікуване вихідне значення $y_{очік}$, точність обчислень ε . Існує можливість оцінки параметрів складності і часу роботи системи у залежності від вибраної точності. Пропонується функція збереження результатів роботи програми і деяких інших сервісних функцій.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ашманов С. А. Введення в математичну економіку / С. А. Ашманов. – М.: Наука, 1984. – 296 с.
2. Бережная Е. В. Математические методы моделирования экономических систем / Е. В. Бережная, В. И. Бережной. – М.: Финансы и статистика, 2003. – 368 с.
3. Бирман И. Я. Оптимальное программирование / И. Я. Бирман. – М.: Экономика, 1968. – 232 с.
4. Горчаков А. А. Компьютерные экономико-математические модели / А. А. Горчаков, И. В. Орлова. – М.: Компьютер, ЮНИТИ, 1995. – 136 с.
5. Жданов С. Экономические модели и методы управления / С. Жданов. – М.: Эльта, 1998. – 102 с.
6. Замков О. О. Математические методы в экономике / О. О. Замков, А. В. Толстонаяченко, Ю. Н. Черемных. – М.: ДНСС, 1997. – 254 с.
7. Ефимова М. Р. Общая теория статистики: учебник / М. Р. Ефимова, Е. В. Петрова, В. Н. Румянцев. – М.: ИНФРА-М, 1998. – 416 с.
8. Интрилигатор М. Математические методы оптимизации и экономическая теория / М. Интрилигатор. – М.: Прогресс, 1975. – 606 с.
9. Карасев А. И. Математические методы и модели в планировании / А. И. Карасев, Н. Ш. Кремер, Т. Н. Савельева. – М. Экономика, 1987. – 356 с.
10. Конюховский П. Математические методы исследования в экономике / П. Конюховский. – СПб.: Питер, 2000. – 208 с.
11. Лагоша Б. А. Оптимальное управление в экономике / Б. А. Лагоша. – М.: Финансы и статистика, 2008. – 224 с.
12. Петров Е. Г. Методи і засоби прийняття рішень у соціально – економічних системах: навч. посіб. / Е. Г. Петров, М. В. Новожилова / За ред. Е. Г. Петрова. – К.: Техніка, 2004. – 256с.
13. Сивый В. Б. Математические методы и модели в планировании и управлении жилищно-коммунальным хозяйством: учеб. пособ. для вузов / В. Б. Сивый, Б. Г. Скоков. – Х.: Основа, 1991. – 208 с.
14. Скурихин В. Н. Математическое моделирование / В. Н. Скурихин, В. Б. Шифрин, В. В. Дубровский. – К.: Техніка, 1983. – 270 с.
15. Сытник В. Ф. Математические модели в планировании и управлении предприятиями / В. Ф. Сытник, Е. А. Каратодава. – К.: Вища школа, 1985. – 214 с.

16. Хазанова Л. Математическое моделирование в экономике / Л. Хазанова. – М.: БЕК, 1998. – 141 с.
17. Ястремський О. І. Основи теорії економічного ризику: навч. посіб. для студентів економічних спеціальностей ВНЗ / О. І. Ястремський. – К.: АртЕк, 1997. – 248 с.
18. Банди Б. Основы линейного программирования / Б. Банди. – М.: Радио и связь, 1989. – 176 с.
19. Дрейпер Н. Прикладной регрессионный анализ / Н. Дрейпер, Г. Смит. – М.: Статистика, 1973. – 392 с.
20. Ларичев О. И. Теория и методы принятия решений / О. И. Ларичев. – М.: Логос, 2000. – 296 с.
21. Ляшенко І. М. Економіко-математичні методи та моделі сталого розвитку / І. М. Ляшенко. – К.: Вища школа, 1999. – 234 с.
22. Малыхин В. И. Математическое моделирование экономики / В. И. Малыхин. – М.: УРАО, 1998. – 423 с.
23. Методические указания к самостоятельному изучению курса «Экономико-математические методы и модели в планировании и управлении», проведению практических занятий и выполнению контрольных работ (для студентов 4, 5 курсов всех форм обучения, специальности 1722) / Составитель Скоков Б. Г. – Х.: Харьковское межвузовское полиграфическое предприятие, 1988. – 58 с.
24. Методична розробка практичного заняття із студентами 4 – 5 курсів з теми: «Оцінка достовірності результатів дослідження» / Укл.: Таралло В. Л., Зубович А. П., Ясинська Е. Ц. – Чернівці, 2001. – 6 с.
25. Миксюк С. Ф. Экономико-математические методы и модели / С. Ф. Миксюк, В. Н. Комкова. – М.: БГЭУ, 2006. – 220 с.
26. Мішура Ю. С. Теоретично-ймовірнісні та статистичні методи в економетриці та фінансовій математиці / Ю. С. Мішура, В. М. Пархоменко, М. Й. Ядренко. – К. Інформтехніка, 1995. – 380 с.
27. Монахов А. Математические методы анализа экономики / А. Монахов. – СПб.: Питер, 2002. – 176 с.
28. Рао С. Р. Линейные статистические методы и их применение / С. Р. Рао. – М.: Наука, 1968. – 548 с.
29. Самойленко М. І. Дослідження операцій (Математичне програмування. Теорія масового обслуговування): навч. посіб. / М. І. Самойленко, Б. Г. Скоков. – Харків: ХНАМГ, 2005. – 176 с.

30. Схрейвер А. Теория линейного и целочисленного программирования / А. Схрейвер. – К. Вища школа, 1991. – 360 с.
31. Терехов Л. Л. Экономико-математические методы / Л. Л. Терехов. – М.: Статистика, 1988. – 166 с.
32. Федосеев В. В. Экономико-математические методы и модели в маркетинге / В. В. Федосеев. – М.: Финстатинформ, 1996. – 110 с.
33. Четыркин Е. М. Статистические методы прогнозирования / Е. М. Четыркин. – М.: Финансы и статистика, 1979. – 302 с.
34. Химмельблау Д. Анализ процессов статистическими методами / Д. Химмельблау. – М.: Наука, 1978. – С. 430-433.
35. Ху Т. Целочисленное программирование и потоки в сетях / Т. Ху. – М.: Мир, 1974. – 520 с.
36. Шелобаев С. И. Математические методы и модели в экономике, финансах, бизнесе: учеб. пособ. для вузов / С. И. Шелобаев. – М.: ЮНИТИ-ДАНА, 2000. – 376 с.
37. Шрейдер Ю. А. Системы и модели / Ю. А. Шрейдер, А. А. Шаров. – М.: Радио и Связь, 1982. – 152 с.
38. Экономико-математические методы и прикладные модели: учеб. пособ. для вузов / В. В. Федосеев, А. Н. Гармаш, Д. М. Дайитбегов и др. – М.: ЮНИТИ, 1999. – 391 с.
39. Hayes R. Dynamic Manufacturing: Creating Learning Organization / R. Hayes, S. C. Wheelwright, K. B. Clark. – The Free Press, NY, 1988. – P. 13-65.
40. Merril William C. Introduction to Economic statistics / C. Merril William, A. Fox Karl. – John: Wiley&Sans, 1970. – 658 p.
41. Альгин А. П. Грани экономического риска / А. П. Альгин. – М.: Знание, 1991. – 64с.
42. Балабанов И. Т. Риск-менеджмент / И. Т. Балабанов. – М.: Финансы и статистика, 1996. – 192 с.
43. Буянов В. П. Рискология (управление рисками) / В. П. Буянов, К. А. Кирсанов, Л. А. Михайлов. – М.: Экзамен, 2003. – 384 с.
44. Вітлінський В. В. Ризик у менеджменті / В. В. Вітлінський, С. І. Наконечний. – К.: Борисфен-М, 1996. – 336 с.
45. Воробьёв Ю. Л. Управление риском и устойчивое развитие. Человеческое измерение / Ю. Л. Воробьёв, Г. Г. Малинецкий, Н. А. Махутов // Общественные науки и современность. – 2000. – № 6. – С. 3-16.
46. Демченков В. С. Системный анализ деятельности предприятий / В. С. Демченков, В. И. Милета. – М.: Финансы и статистика, 1990. – 182 с.

47. Дубров А. М. Моделирование рисков ситуации в экономике и бизнесе: учеб. пособ. / А. М. Дубров, Б. А. Лагоша, Е. Ю. Хрусталеv; под ред. Б. А. Лагоши. – М.: Финансы и статистика, 2000. – 176 с.
48. Лапуста М. Г. Риски в предпринимательской деятельности / М. Г. Лапуста, Л. Г. Шаршукова. – М.: Инфра-М, 1998. – 225 с.
49. Макаревич Л. М. Управление предпринимательскими рисками / Л. М. Макаревич. – М.: Дело и Сервис, 2006. – 448 с.
50. Малинецкий Г. Г. Управление риском и редкие катастрофические события / Г. Г. Малинецкий // Математическое моделирование. – 2002. – Т. 14. – № 8. – С. 107-112.
51. Методичні вказівки «Використання пакету програм «Statistica» в економетричних дослідженнях» (для студентів 3 курсу денної форми навчання, спец. 6.050200 «Менеджмент організацій») / Укл.: Скоков Б. Г., Мамонов К. А. – Х.: ХНАМГ, 2007. – 51 с.
52. Ракитов А. И. Принципы научного мышления / А. И. Ракитов. – М.: Политиздат, 1975. – 143 с.
53. Райзберг Б. А. Предпринимательство и риск / Б. А. Райзберг // Новое в жизни, науке и технике. – 1992. – № 4. – 61 с.
54. Риски в современном бизнесе / П. Г. Грабовый, С. Н. Петрова, С. И. Полтавцев и др. – М.: Алане, 1994. – 200 с.
55. Сергеев М. Предпринимательский риск и стратегии предпринимателя [Электронный ресурс]. – Режим доступа к сайту: <http://www.fact.ru/archiv/num01/serg.html>.
56. Тони Райс Финансовые инвестиции и риск: пер. с англ. / Тони Райс, Брайан Койли. – Торгово-издательское бюро BVH, 1995. – 592 с.
57. Уткин Э. А. Риск-менеджмент: учебник / Э. А. Уткин. – М.: Тандем, 1998. – 288 с.
58. Чернов В. А. Анализ коммерческого риска / В. А. Чернов. – М.: Финансы и статистика, 1998. – 128 с.
59. Хохлов Н. В. Управление риском: учеб. пособ. для вузов / Н. В. Хохлов. – М.: ЮНИТИ-ДАНА, 1999. – 239 с.
60. Ястремський О. І. Моделювання економічного ризику / О. І. Ястремський. – К.: Либідь, 1992. – 176 с.
61. Daenzer B. J. Fact-Finding Techniques in Risk Analysis / B. J. Daenzer. – AMA, 1970. – P. 63-67.
62. Head G. Essentials of Risk Management / G. Head, S. Horn. – John: Wiley&Sans, 1991. – V. – P. 136.

63. Robert N. Charette. Applications Strategies for Risk Analysis / Robert N. Charette // McGraw-Hill Book Company. – 1990. – ISBN 0-07-010888-9.
64. Simon J. D. Political Risk Assessment / Simon J. D. // Columbia Journal of World Business. – 1982. – V 17, № 3. – P. 146-157.
65. Грубер Й. Економетрія: посібник / Й. Грубер. – К.: ЗАТ «Нічлава», 1998. – 295 с.
66. Джонстон Д. Ж. Эконометрические методы / Д. Ж. Джонстон. – М.: Финансы и статистика, 1980. – 277 с.
67. Доугерти К. Введение в эконометрику: пер. с англ. / К. Доугерти. – М.: ИНФРА-М, 2001. – 402 с.
68. Лещинський О. Л. Економетрія: навч. посіб. для студ. ВНЗ / О. Л. Лещинський, В. В. Рязанцева, О. О. Юнькова. – Л.: МАУП, 2003. – 208 с.
69. Лук'яненко І. Г. Економетрика: підручник / І. Г. Лук'яненко, Л. І. Краснікова. – К.: Знання, КОО, 1998. – 494 с.
70. Методичні вказівки до виконання практичних завдань і самостійної роботи з дисципліни «Економетрія» (для студентів 3 курсу денної форми навчання спец. 7.050201 «Менеджмент організацій») / Укл.: Мамонов К. А. – Х.: ХНАМГ, 2006. – 27 с.
71. Робоча програма і короткий конспект лекцій до самостійного вивчення курсу «Економетрія» (для студентів денної і заочної форм навчання спеціальностей «Менеджмент організацій», «Облік і аудит» та «Економіка підприємства») / Укл.: Скоков Б. Г., Мамонов К. А. – Х.: ХНАМГ, 2006. – 105 с.
72. Lofti V. Decision Support System for Production and Operations Managament (DSSPOW) / V. Lofti, C. Pegels. – IRWIN, 1991. – 359 с.
73. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений / Л. Заде. – М.: Мир, 1976. – 166 с.
74. Tsoukalas L. H. Fuzzy and Neural Approaches in Engineering / L. H. Tsoukalas, R. E. Uhrig. – New York: John Wiley&Sons.Inc, 1997. – 587 p.
75. Круглов В. В. Самоорганизующаяся система нечеткого логического вывода / В. В. Круглов, Н. П. Прохоренкова // Перспективные информационные технологии и интеллектуальные системы. – 2004. – № 3 (19). – С. 4-12.
76. Сайт Международной ассоциации нечетких систем «International Fuzzy Systems Association» [Электронный ресурс]. – Режим доступа к сайту: <http://www.abo.fi/~rfuller/ifsa.html>.
77. Сайт «Нечеткая логика, нечеткие системы и мягкие вычисления» [Электронный ресурс]. – Режим доступа к сайту: <http://www.kstu.ru/P/fuzzy/>.

78. Свободная энциклопедия [Электронный ресурс]. – Режим доступа к энциклопедии: <http://ru.wikipedia.org/>.

79. Кучеренко Є. І. Інтелектуальні технології моделювання та аналізу взаємодіючих процесів розширеними інтерпретованими нечіткими мережами Петрі / Є. І. Кучеренко // Вісник Технологічного університету Поділля. – 2002. – Т. 1. – Вип. 3 (41). – С. 118-122.

80. Кучеренко Е. И. Прикладные аспекты моделирования нечетких процессов в сложных системах / Е. И. Кучеренко, И. С. Творошенко // Сборник научных трудов ХУВС. – 2010. – Вып. 1 (23). – С. 127-131.

81. Бодянский Е. В. О синтезе нечетких алгоритмов на основе композиции фрагментов правил и моделей / Е. В. Бодянский, Е. И. Кучеренко, И. С. Творошенко // АСУ и приборы автоматики. – 2004. – Вып. 128. – С. 19-28.

82. Кучеренко Е. И. Интеллектуальные технологии в задачах принятия решений технологических комплексов на основе нечеткой интервальной логики / Е. И. Кучеренко, В. А. Филатов, И. С. Творошенко, Р. Н. Байдан // Восточно-Европейский журнал передовых технологий. – 2005. – № 2. – С. 92-96.

83. Журкин И. Г. Геоинформационные системы / И. Г. Журкин, С. В. Шайтура. – М.: Кудиц-Пресс, 2009. – 272 с.

84. Градиентные методы и оптимизация [Электронный ресурс]. – Режим доступа к источнику: http://opds.sut.ru/electronic_manuals/pe/f022.htm.

85. Кучеренко Е. И. О методах настройки функций принадлежности в нечетких системах / Е. И. Кучеренко, А. В. Корниловский, И. С. Творошенко // Системы управления, навигации и связи. – 2010. – Вып. 1 (13). – С. 94-98.

86. Левитин А. В. Алгоритмы. Введение в разработку и анализ / А. В. Левитин. – М.: Вильямс, 2006. – 576 с.

87. Шилдт Г. Полный справочник по C++, 4-е издание / Г. Шилдт. – М.: Вильямс, 2004. – 800 с.

88. Лапчик М. П. Численные методы. Серия Высшее профессиональное образование. Информатика и вычислительная техника / М. П. Лапчик, М. И. Рагулина, Е. К. Хеннер. – М.: Академия, 2007. – 384 с.

Навчальне видання

МАМОНОВ Костянтин Анатолійович

ТВОРОШЕНКО Ірина Сергіївна

МАТЕМАТИЧНІ МЕТОДИ І МОДЕЛІ В ОЦІНЦІ НЕРУХОМОСТІ

НАВЧАЛЬНИЙ ПОСІБНИК

Відповідальний за випуск: *К. А. Мамонов*

За авторською редакцією

Комп'ютерний набір: *І. С. Творошенко*

Комп'ютерне верстання: *І. В. Волосожарова*

Дизайн обкладинки: *Д. В. Шаульський*

Підп. до друку 26.06.2014 р.

Друк на різнографі.

Тираж 300 пр.

Формат 60x84/16

Ум. друк. арк. 12,5

Зам. №

Видавець і виготовлювач:

Харківський національний університет
міського господарства імені О. М. Бекетова,
вул. Революції, 12, Харків, 61002

Електронна адреса: rectorat@kname.edu.ua

Свідоцтво суб'єкта видавничої справи:

ДК № 4705 від 28.03.2014 р.